FIGURE 1A

```
CCCGTGCCCC TAAAGGCCGC CGAGAAAGCT AAGTCCAAAT GTGACGTCGG 50
AGGTCTCGAC ATGGTCGCCA ACCCTCCAAA TGCTACCCGC CGGCCCACGC 100
AACGCGGGCT TTTATAAAGA TGGCGCGCGA GACAATAACA CTTACTCATC 150
CGCGTACGCG TTTATTATTG TCAATATTTG TGTGGTTATT ATTACTGCTA 200
CCGCCCTTGT TTCTGCAAGG CCCTCGCCGC GGCCCAGGCC ACTATTCCGG 250
CAGCGGCCGC CGACGCGGCG AGCGTCGCCG CTAACGTCGG CGCCGCGGGG 300
AGCGGGGTTT CTTCGACTTA AATAGACTCC CGAGAAAAAA TTTTGGCTGC 350
CGTTCGCCAT CATCCGAGTC GGAAACACAG TATGCGGCCG AGTTAGGTTT 400
TACTTTTAAA AACTTTACCG TGCTGTACGG CCAGGGCGTT CTCAGGCTCG 450
AAGGGGCAAG AGTTGTCCAG ACTGATGGGT GACTCAGAGA CAGCGTTGTC 500
TTGTCTCCGT TTACCAAAAA TATTTCCACT CCTCTCTCAA AATTTTTACC 550
TCCGGTTTCG GTAATTAGGA AAGTTTTTGG CGCAGGGAGG TTTAAAGCTG 600
CCATGCATAT GTCAGCGGTA CCCAGCACCC ACAAATGGAA CTCTTTTGCG 650
GCATACGCGC CAGATGACAA ATGGTAAAAC CCTGCGTCCA AGCCGCTCCA 700
CTCGGGACTT ACTCCAGGCG GGTCGCCCCC CTCACCGAAC CGAATCACGG 750
GTCTGCACAT CCTGGGAAGG GAAAACAGCT CCCCGGAAAC TTCGTACAGA 800
GATGCCGGGC GCACGATTAC CGATAATGTA CTCGGACGAT CGTAACTCGC 850
CATAGTTTTC ACTGCGTGAA CCAATTCTTT CCATCCAGAA TCCGAGAGCT 900
CAAATCTAGA ATTAGGTAGT TTGTAGTGCG AATCGACCGC AGAAACTATA 950
GTCACTTTTA CAGGCGCCAT CGCCGCTCAG ACTCCACCCC GCTATGATGT 1000
CAGAAATATA ACGCTCTTAT TCTAGCAGAG TCAGGCCAAT ATATACAGCT 1050
TAGAGAAGAT GCGGTTTCGG CGCATCTGTT CACGCTCTAG GGCAGAAAAA 1100
CGAAGAAGAA CAACCGAGAA TCCGCTTACC TCAAAACGCG TTTGCGTATT 1150
GGATAGTTTC TCACGGACAA TGTCATTGCG CCCCTATGCA GAAATTTTGC 1200
CGACCGCGGA AGGCGTCGAG CGCCTCGCCG AACTTGTTAG TGTGACAATG 1250
ACAGAACGCG CGGAACCTGT GACAGAGAAT ACAGCTGTAA ACAGTATCCC 1300
CCCGGCTAAC GAGAACGGGC AGAACTTCGC ATATGCAGGC GATGGGCCCT 1350
CGACTACTGA AAAAGTTGAC GGCTCGCATA CAGACTTCGA TGAAGCATCG 1400
AGCGACTACG CCGGCCCTGT CCCGCTCGCG CAAACTAGAT TGAAGCATTC 1450
GGATGAATTT CTTCAGCACT TCCGAGTTTT AGACGATTTG GTGGAGGGGG 1500
CTTACGGGTT TATCTGCGGC GTCCGTCGCT ACACCGAGGA AGAGCAACGT 1550
CGAAGAGGGG TTAACAGTAC TAACCAGGGG AAATCAAAAT GTAAGCGCCT 1600
GATAGCTAAA TATGTGAAAA ATGGAACAAG GGCGGCCTCT CAGCTGGAAA 1650
ATGAAATTTT GGTTCTCGGG CGCCTAAATC ACGAGAATGT TCTCAAGATC 1700
CAGGAAATCC TTCGGTACCC GGATAATACG TACATGTTAA CGCAGAGGTA 1750
```

FIGURE 1B

```
TCAGTTCGAC TTGTACAGCT ACATGTACGA TGAAGCGTTC GACTGGAAAG 1800
ACAGTCCAAT GCTTAAACAG ACTAGACGCA TCATGAAGCA GCTCATGTCA 1850
GCGGTCTCGT ATATCCATTC AAAGAAACTG ATTCACAGGG ACATCAAACT 1900
CGAAAATATT TTCTTAAACT GCGACGGCAA GACAGTGCTG GGCGACTTTG 1950
GAACTGTCAC GCCTTTTGAA AATGAGCGGG AGCCCTTCGA ATATGGATGG 2000
GTGGGGACCG TGGCTACTAA CTCTCCCGAG ATACTCGCCA GGGATTCGTA 2050
CTGTGAAATT ACAGACATTT GGAGCTGCGG AGTAGTATTG CTGGAAATGG 2100
TAAGCCATGA ATTTTGCCCG ATCGGCGATG GCGGGGGAAA TCCGCACCAG 2150
CAATTGCTGA AAGTTATCGA CTCTCTCTCA GTTTGTGATG AAGAGTTCCC 2200
AGACCCCCCG TGTAATCTGT ACAATTATTT GCATTATGCG AGCATCGATC 2250
GCGCCGGACA TACGGTCCCG TCGCTCATAC GGAACCTCCA CCTTCCGGCG 2300
GATGTGGAAT ACCCTCTAGT TAAAATGCTT ACTTTTGACT GGCGTTTGAG 2350
ACCCAGCGCG GCCGAAGTAT TGGCAATGCC ACTGTTTTCG GCTGAAGAGG 2400
AACGGACCAT AACAATTATT CATGGAAAAC ATAAACCCAT CCGACCCGAA 2450
ATCCGTGCGC GGGTGCCACG GTCCATGAGT GAAGGTTAAT AATAAAGGAC 2500
GGAGATAGAG AACTGAAGCG TCAGATTTTT TTAAAAAAAT AAATGATCGA 2550
GAACTTATGA TTTGTCTTTC TTGAATGACC TTGCCCCATC GATTAACGAA 2600
AAGACCTTTC GCGCGTCGAT TCTGCTCGGT CTTTGTGATA CATTATAGTG 2650
AGACTAAACT CGACCGATAT AACAAGACAA TGTTACTCTA TAGACCGGAC 2700
TCAACCATGC GGCATAGCGG AGGCGACGCA AATCACAGAG GGATAAGGCC 2750
GAGGCGGAAA TCTATTGGAG CGTTTAGCGC GCGCGAAAAG ACTGGAAAAC 2800
GAAATGCGCT GACGGAAAGC AGCTCCTCCT CCGACATGCT AGATCCGTTT 2850
TCCACGGATA AGGAATTTGG CGGTAAGTGG ACGGTAGACG GACCTGCCGA 2900
CATTACTGCC GAGGTCCTTT CTCAGGCATG GGACGTTCTC CAATTAGTGA 2950
AGCATGAAGA TGCGGAGGAG GAGAGAGTGA CTTATGAGTC CAAACCGACC 3000
CCGATACAGC CGTTCAATGC CTGGCCGGAC GGGCCGAGTT GGAACGCGCA 3050
GGATTTTACT CGAGCGCCAA TAGTTTATCC CTCTGCGGAG GTATTGGACG 3100
CAGAGGCGTT GAAAGTAGGG GCATTCGTTA GCCGAGTTTT ACAATGTGTA 3150
CCGTTCACGC GATCAAAGAA AAGCGTTACG GTGCGGGATG CGCAGTCGTT 3200
TTTGGGGGAC TCGTTCTGGA GAATAATGCA GAACGTTTAC ACGGTTTGCT 3250
TACGACAGCA CATAACTCGA CTCAGGCACC CTTCCAGCAA AAGCATTGTT 3300
AACTGCAACG ACCCTCTATG GTACGCCTAC GCGAATCAAT TTCACTGGAG 3350
AGGAATGCGC GTGCCGTCGC TTAAATTAGC CTCTCCCCCG GAGGAGAATA 3400
TTCAACACGG CCCAATGGCC GCCGTTTTTA GAAACGCGGG GGCTGGTCTG 3450
TTCCTGTGGC CTGCCATGCG CGCAGCCTTT GAAGAGCGCG ACAAGCGACT 3500
```

FIGURE 1C

```
GTTAAGAGCA TGCCTGTCTT CACTCGATAT CATGGACGCA GCCGTCCTCG 3550
CGTCGTTTCC ATTTTACTGG CGCGGCGTCC AAGACACCTC GCGCTTCGAG 3600
CCTGCGCTGG GCTGTTTGTC AGAGTACTTT GCACTAGTGG TGTTACTGGC 3650
CGAGACGGTC TTAGCGACCA TGTTCGACCA CGCACTGGTA TTCATGAGGG 3700
CGCTGGCAGA CGGCAATTTC GATGACTATG ACGAAACTAG ATATATAGAC 3750
CCCGTTAAAA ACGAGTACCT GAACGGAGCC GAGGGTACTC TGTTACGGGG 3800
CATAGTGGCC TCCAACACCG CTCTGGCGGT GGTTTGCGCA AACACCTATT 3850
CGACGATAAG AAAACTCCCG TCCGTGGCAA CTAGCGCGTG CAATGTTGCC 3900
TACAGGACCG AAACGCTGAA AGCGAGGCGC CCTGGCATGA GCGACATATA 3950
CCGGATATTA CAAAAAGAGT TTTTCTTTTA CATTGCGTGG CTCCAGAGGG 4000
TTGCAACACA CGCAAATTTC TGTTTAAACA TTCTGAAGAG AAGCGTGGAT 4050
ACGGGCCCCC GCCATTTTTG TTCAGGGCCA GCTCGGAGAA GCGGCTGCAG 4100
CAGTTAAATA AAATGCTCTG CCCCCTTCTC GTGCCGATTC AATATGAAGA 4150
CTTTTCGAAG GCCATGGGGT CTGAGCTCAA GAGGGAAAAG TTAGAGACAT 4200
TCGTTAAAGC TATTTCCAGC GACAGGGACC CGAGGGGGTC CTTAAGATTT 4250
CTCATTTCGG ACCATGCAAG GGAAATTATT GCAGACGGAG TACGGTTTAA 4300
GCCGGTGATA GACGAGCCGG TTCGGGCTTC AGTTGCGCTG AGTACCGCTG 4350
CCGCTGGGAA AGTGAAAGCG CGACGCTTAA CCTCAGTTCG CGCGCCCGTA 4400
CCGCCCGCAG GCGCCGTTTC CGCGCGCCGG AAATCGGAAA TATGATAAAA 4450
ATGCTTGGCA TTTGCGGGCG AAGAGGCGTG ATCTGAAGGG CTCCACAATG 4500
ACGTAACTGA GCTACGCATC CCTATAAAGT GTACSCGCTG ACCGCTAGCC 4550
CATACAGTGT TACAGGAGGG GAGAGAGACA ACTTCAGCTC GAAGTCTGAA 4600
GAGACATCAT GAGCGGCTTC AGTAACATAG GATCGATTGC CACCGTTTCC 4650
CTAGTATGCT CGCTTTTGTG CGCATCTGTA TTAGGGGCGC CGGTACTGGA 4700
CGGGCTCGAG TCGAGCCCTT TCCCGTTCGG GGGCAAAATT ATAGCCCAGG 4750
CGTGCAACCG CACCACGATT GAGGTGACGG TCCCGTGGAG CGACTACTCT 4800
GGTCGCACCG AAGGAGTGTC AGTCGAGGTG AAATGGTTCT ACGGGAATAG 4850
TAATCCCGAA AGCTTCGTGT TCGGGGTGGA TAGCGAAACG GGCAGTGGAC 4900
ACGAGGACCT GTCTACGTGC TGGGCTCTAA TCCATAATCT GAACGCGTCT 4950
GTGTGCAGGG CGTCTGACGC CGGGATACCT GATTTCGACA AGCAGTGCGA 5000
AAAAGTGCAG AGAAGACTGC GCTCCGGGGT GGAACTTGGT AGTTACGTGT 5050
CTGGCAATGG ATCCCTGGTG CTGTACCCAG GGATGTACGA TGCCGGCATC 5100
TACGCCTACC AGCTCTCAGT GGGTGGGAAG GGATATACCG GGTCTGTTTA 5150
TCTAGACGTC GGACCAAACC CCGGATGCCA CGACCAGTAT GGGTACACCT 5200
ATTACAGCCT GGCCGACGAG GCGTCAGACT TATCATCTTA TGACGTAGCC 5250
```

FIGURE 1D

```
TCGCCCGAAC TCGACGGTCC TATGGAGGAA GATTATTCCA ATTGTCTAGA 5300
CATGCCCCCG CTACGCCCAT GGACAACCGT TTGTTCGCAT GACGTCGAGG 5350
AGCAGGAAAA CGCCACGGAC GAGCTTTACC TATGGGACGA GGAATGCGCC 5400
GGTCCGCTGG ACGAGTACGT CGACGAAAGG TCAGAGACGA TGCCCAGGAT 5450
GGTTGTCTTT TCACCGCCCT CTACGCTCCA GCAGTAGCCA CCCGAGAGTG 5500
TTTTTTGTGA GCGCCCACGC AACATACCTA ACTGCTTCAT TTCTGATCAA 5550
TTATTGCGTA TTGAATAAAT AAACAGTACA AAAGCATCAG GTGTGGTTTG 5600
CGTGTCTGTG CTAAACCATG GCGTGTGCGG GTGAAAcCGT AAATTACGTG 5650
ATAATAAATA GCATAGGAGT TGGCGTGCAG CGTATTTCGC CGAGAGATGG 5700
GGACAATGTT AGTGTTGCGC CTTTTCCTAC TTGCAGTAGC GGACGCGGCG 5750
TTGCCGACCG GCAGATTCTG CCGAGTTTGG AAGGTGCCTC CGGGAGGAAC 5800
CATCCAAGAG AACCTGGCGG TGCTCGCGGA ATCGCCGGTC ACGGGACACG 5850
CGACATATCC GCCGCCTGAA GGCGCCGTCA GCTTTCAGAT TTTTGCGGAC 5900
ACCCCTACTT TGCGCATTCG CTACGGGCCT ACGGAGGACG AACTTGCACT 5950
GGAGCGCGGG ACGTCCGCCT CAGACGCGGA CAACGTGACA TTTTCGCTGT 6000
CATATCGCCC GCGCCCAGAA ATTCACGGAG CATACTTCAC CATAGGGGTA 6050
TTCGCTACTG GCCAGAGCAC GGAAAGCAGC TATTCGGTCA TCAGTCGGGT 6100
CTTAGTTAAC GCCTCTCTGG AACGGTCCGT GCGCCTGGAA ACGCCGTGCG 6150
ATGAAAATTT TTTGCAGAAC GAGCCTACAT GGGGCTCGAA GCGTTGGTTA 6200
GGCCCCCCGT CGCCTTATGT GCGAGATAAC GATGTCGCCG TGTTGACAAA 6250
AGCGCAGTAC ATTGGGGAGT GCTACTCCAA CTCGGCGGCC CAGACGGGGC 6300
TCACGTCTCT CAACATGACC TTTTTCTATT CGCCTAAAAG AATAGTAAAC 6350
GTCACGTGGA CAACCGGCGG CCCCTCCCCC TCGCGCATAA CGGTATACTC 6400
GTCGCGGGAG AACGGGCAGC CCGTGTTGAG GAACGTTTCT GACGGGTTCT 6450
TGGTTAAGTA CACTCCCGAC ATTGACGGCC GGGCCATGAT AAACGTTATT 6500
GCCAATTATT CGCCGGCGGA CTCCGGCAGC GTCCTCGCGT TTACGGCCTT 6550
TAGGGAAGGA AAACTCCCAT CCGCGATTCA ACTGCACCGG ATAGATATGT 6600
CCGGGACTGA GCCGCCGGGG ACTGAAACGA CCTTCGACTG TCAAAAAATG 6650
ATAGAAACCC CGTACCGAGC GCTCGGGAGC AATGTTCCCA GGGACGACTC 6700
TATCCGTCCG GGGGCCACTC TGCCTCCGTT CGATACCGCA GCACCTGATT 6750
TCGATACAGG TACTTCCCCG ACCCCCACTA CCGTGCCAGA GCCAGCCATT 6800
ACTACACTCA TACCGCGCAG CACTAGCGAT ATGGGATTCT TCTCCACGGC 6850
ACGTGCTACC GGATCAGAAA CTCTTTCGGT ACCCGTCCAG GAAACGGATA 6900
GAACTCTTTC GACAACTCCT CTTACCCTTC CACTGACTCC CGGTGAGTCA 6950
GAAAATACAC TGTTTCCTAC GACCGCGCCG GGGATTTCTA CCGAGACCCC 7000
```

FIGURE 1E

```
GAGCGCGGCA CATGAAACTA CACAGACCCA GAGTGCAGAA ACGGTGGTCT 7050
TTACTCAGAG TCCGAGTACC GAGTCGGAAA CCGCGCGGTC CCAGAGTCAG 7100
GAACCGTGGT ATTTTACTCA GACTCCGAGT ACTGAACAGG CGGCTCTTAC 7150
TCAGACGCAG ATCGCAGAAA CGGAGGCGTT GTTTACTCAG ACTCCGAGTG 7200
CTGAACAGAT GACTTTTACT CAGACTCCGG GTGCAGAAAC CGAGGCACCT 7250
GCCCAGACCC CGAGCACGAT ACCCGAGATA TTTACTCAGT CTCGTAGCAC 7300
GCCCCCCGAA ACCGCTCGCG CTCCGAGCGC GGCGCCGGAG GTTTTTACAC 7350
AGAGTTCGAG TACGGTAACG GAGGTGTTTA CTCAGACCCC GAGCACGGTA 7400
CCGAAAACTA CTCTGAGTTC GAGTACTGAA CCGGCGATTT TTACTCGGAC 7450
TCAGAGCGCG GGAACTGAGG CCTTTACTCA GACTTCGAGT GCCGAGCCGG 7500
ACACTATGCG AACTCAGAGT ACTGAAACAC ACTTTTTCAC TCAGGCCCCG 7550
AGTACGGTAC CGAAAGCTAC TCAGACTCCG AGTACAGAGC CGGAGGTGTT 7600
GACTCAGAGT CCGAGTACCG AACCTGTGCC TTTCACCCGG ACTCTGGGCG 7650
CAGAGCCGGA AATTACTCAG ACCCCGAGCG CGGCACCGGA GGTTTATACT 7700
CGGAGTTCGA GTACGATGCC AGAAACTGCA CAGAGCACAC CCCTGGCCTC 7750
GCAAAACCCT ACCAGTTCGG GAACCGGGAC GCATAATACT GAACCGAGGA 7800
CTTATCCAGT GCAAACGACA CCACATACCC AGAAACTCTA CACAGAAAAT 7850
AAGACTTTAT CGTTTCCTAC TGTTGTTTCA GAATTCCATG AGATGTCGAC 7900
GGCAGAGTCG CAGACGCCCC TATTGGACGT CAAAATTGTA GAGGTGAAGT 7950
TTTCAAACGA TGGCGAAGTA ACGGCGACTT GCGTTTCCAC CGTCAAATCT 8000
CCCTATAGGG TAGAAACTAA TTGGAAAGTA GACCTCGTAG ATGTAATGGA 8050
TGAAATTTCT GGGAACAGTC CCGCCGGGGT TTTTAACAGT AATGAGAAAT 8100
GGCAGAAACA GCTGTACTAC AGAGTAACCG ATGGAAGAAC ATCGGTCCAG 8150
CTAATGTGCC TGTCGTGCAC GAGCCATTCT CCGGAACCTT ACTGTCTTTT 8200
CGACACGTCT CTTATAGCGA GGGAAAAAGA TATCGCGCCA GAGTTATACT 8250
TTACCTCTGA TCCGCAAACG GCATACTGCA CAATAACTCT GCCGTCCGGC 8300
GTTGTTCCGA GATTCGAATG GAGCCTTAAT AATGTTTCAC TGCCGGAATA 8350
TTTGACGGCC ACGACCGTTG TTTCGCATAC CGCTGGCCAA AGTACAGTGT 8400
GGAAGAGCAG CGCGAGAGCA GGCGAGGCGT GGATTTCTGG CCGGGGAGGC 8450
AATATATACG AATGCACCGT CCTCATCTCA GACGGCACTC GCGTTACTAC 8500
GCGAAAGGAG AGGTGCTTAA CAAACACATG GATTGCGGTG GAAAACGGTG 8550
CTGCTCAGGC GCAGCTGTAT TCACTCTTTT CTGGACTTGT GTCAGGATTA 8600
TGCGGGAGCA TATCTGCTTT GTACGCAACG CTATGGACCG CCATTTATTT 8650
TTGAGGAATG CTTTTTGGAC TATCGTACTG CTTTCTTCCT TCGCTAGCCA 8700
GAGCACCGCC GCCGTCACGT ACGACTACAT TTTAGGCCGT CGCGCGCTCG 8750
```

FIGURE 1F

```
ACGCGCTAAC CATACCGGCG GTTGGCCCGT ATAACAGATA CCTCACTAGG 8800
GTATCAAGAG GCTGCGACGT TGTCGAGCTC AACCCGATTT CTAACGTGGA 8850
CGACATGATA TCGGCGGCCA AAGAAAAAGA GAAGGGGGGC CCTTTCGAGG 8900
CCTCCGTCGT CTGGTTCTAC GTGATTAAGG GCGACGACGG CGAGGACAAG 8950
TACTGTCCAA TCTATAGAAA AGAGTACAGG GAATGTGGCG ACGTACAACT 9000
GCTATCTGAA TGCGCCGTTC AATCTGCACA GATGTGGGCA GTGGACTATG 9050
TTCCTAGCAC CCTTGTATCG CGAAATGGCG CGGGACTGAC TATATTCTCC 9100
CCCACTGCTG CGCTCTCTGG CCAATACTTG CTGACCCTGA AAATCGGGAG 9150
ATTTGCGCAA ACAGCTCTCG TAACTCTAGA AGTTAACGAT CGCTGTTTAA 9200
AGATCGGGTC GCAGCTTAAC TTTTTACCGT CGAAATGCTG GACAACAGAA 9250
CAGTATCAGA CTGGATTTCA AGGCGAACAC CTTTATCCGA TCGCAGACAC 9300
CAATACACGA CACGCGGACG ACGTATATCG GGGATACGAA GATATTCTGC 9350
AGCGCTGGAA TAATTTGCTG AGGAAAAAGA ATCCTAGCGC GCCAGACCCT 9400
CGTCCAGATA GCGTCCCGCA AGAAATTCCC GCTGTAACCA AGAAAGCGGA 9450
AGGGCGCACC CCGGACGCAG AAAGCAGCGA AAAGAAGGCC CCTCCAGAAG 9500
ACTCGGAGGA CGACATGCAG GCAGAGGCTT CTGGAGAAAA TCCTGCCGCC 9550
CTCCCCGAAG ACGACGAAGT CCCCGAGGAC ACCGAGCACG ATGATCCAAA 9600
CTCGGATCCT GACTATTACA ATGACATGCC CGCCGTGATC CCGGTGGAGG 9650
AGACTACTAA AAGTTCTAAT GCCGTCTCCA TGCCCATATT CGCGGCGTTC 9700
GTAGCCTGCG CGGTCGCGCT CGTGGGGCTA CTGGTTTGGA GCATCGTAAA 9750
ATGCGCGCGT AGCTAATCGA GCCTAGAATA GGTGGTTTCT TCCTACATGC 9800
CACGCCTCAC GCTCATAATA TAAATCACAT GGAATAGCAT ACCAATGCCT 9850
ATTCATTGGG ACGTTCGAAA AGCATGGCAT CGCTACTTGG AACTCTGGCT 9900
CTCCTTGCCG CGACGCTCGC ACCCTTCGGC GCGATGGGAA TCGTGATCAC 9950
TGGAAATCAC GTCTCCGCCA GGATTGACGA CGATCACATC GTGATCGTCG 10000
CGCCTCGCCC CGAAGCTACA ATTCAACTGC AGCTATTTTT CATGCCTGGC 10050
CAGAGACCCC ACAAACCCTA CTCAGGAACC GTCCGCGTCG CGTTTCGGTC 10100
TGATATAACA AACCAGTGCT ACCAGGAACT TAGCGAGGAG CGCTTTGAAA 10150
ATTGCACTCA TCGATCGTCT TCTGTTTTTG TCGGCTGTAA AGTGACCGAG 10200
TACACGTTCT CCGCCTCGAA CAGACTAACC GGACCTCCAC ACCCGTTTAA 10250
GCTCACTATA CGAAATCCTC GTCCGAACGA CAGCGGGATG TTCTACGTAA 10300
TTGTTCGGCT AGACGACACC AAAGAACCCA TTGACGTCTT CGCGATCCAA 10350
CTATCGGTGT ATCAATTCGC GAACACCGCC GCGACTCGCG GACTCTATTC 10400
CAAGGCTTCG TGTCGCACCT TCGGATTACC TACCGTCCAA CTTGAGGCCT 10450
ATCTCAGGAC CGAGGAAAGT TGGCGCAACT GGCAAGCGTA CGTTGCCACG 10500
```

FIGURE 1G

```
GAGGCCACGA CGACCAGCGC CGAGGCGACA ACCCCGACGC CCGTCACTGC 10550
AACCAGCGCC TCCGAACTTG AAGCGGAACA CTTTACCTTT CCCTGGCTAG 10600
AAAATGGCGT GGATCATTAC GAACCGACAC CCGCAAACGA AAATTCAAAC 10650
GTTACTGTCC GTCTCGGGAC AATGAGCCCT ACGCTAATTG GGGTAACCGT 10700
GGCTGCCGTC GTGAGCGCAA CGATCGGCCT CGTCATTGTA ATTTCCATCG 10750
TCACCAGAAA CATGTGCACC CCGCACCGAA AATTAGACAC GGTCTCGCAA 10800
GACGACGAAG AACGTTCCCA AACTAGAAGG GAATCGCGAA AATTTGGACC 10850
CATGGTTGCG TGCGAAATAA ACAAGGGCGC TGACCAGGAT AGTGAACTTG 10900
TGGAACTGGT TGCGATTGTT AACCCGTCTG CGCTAAGCTC GCCCGACTCA 10950
ATAAAAATGT GATTAAGTCT GAATGTGGCT CTCCAATCAT TTCGATTCTC 11000
TAATCTCCCA ATCCTCTCAA AAGGGGCAGT ATCGGACACG GACTGGGAGG 11050
GGCGTACTAC ACGATAGTTA TATGGTACAG CAGAGGCCTC TGAACACTTA 11100
GGAGGAGAAT TCAGCCGGGG AGAGCCCCTG TTGAGTAGGC TTGGGAGCAT 11150
ATTGCAGGAT GAACATGTTA GTGATAGTTC TCGCCTCTTG TCTTGCGCGC 11200
CTAACTTTTG CGACGCGACA CGTCCTCTTT TTGGAAGGCA CTCAGGCTGT 11250
CCTCGGGGAA GATGATCCCA GAAACGTTCC GGAAGGGACT GTAATCAAAT 11300
GGACAAAAGT CCTGCGGAAC GCGTGCAAGA TGAAGGCGGC CGATGTCTGC 11350
TCTTCGCCTA ACTATTGCTT TCATGATTTA ATTTACGACG GAGGAAAGAA 11400
AGACTGCCCG CCCGCGGGAC CCCTGTCTGC AAACCTGGTA ATTTTACTAA 11450
AGCGCGGCGA AAGCTTCGTC GTGCTGGGTT CTGGGCTACA CAACAGCAAT 11500
ATAACTAATA TCATGTGGAC AGAGTACGGA GGCCTGCTCT TTGATCCTGT 11550
AACTCGTTCG GACGAGGGAA TCTATTTTCG ACGGATCTCT CAGCCAGATC 11600
TGGCCATGGA AACTACATCG TACAACGTCA GCGTTCTTTC GCACGTAGAC 11650
GAGAAGGCTC CAGCACCGCA CGAGGTGGAG ATAGACACCA TCAAGCCGTC 11700
AGAGGCCCAC GCGCACGTGG AATTACAAAT GCTGCCGTTT CATGAACTCA 11750
ACGACAACAG CCCCACCTAT GTGACCCCTG TTCTTAGAGT CTTCCCACCG 11800
ACCGAGCACG TAAAATTTAA CGTTACGTAT TCGTGGTATG GGTTTGATGT 11850
CAAAGAGGAG TGCGAAGAAG TGAAACTGTT CGAGCCGTGC GTATACCATC 11900
CTACAGACGG CAAATGTCAG TTTCCCGCAA CCAACCAGAG ATGCCTCATA 11950
GGATCTGTCT TGATGGCGGA ATTCTTGGGC GCGGCCTCTT TGCTGGATTG 12000
TTCCCGCGAT ACTCTAGAAG ACTGCCACGA AAATCGCGTG CCGAACCTAC 12050
GGTTCGATTC GCGACTCTCC GAGTCACGCG CAGGCCTGGT GATCAGTCCT 12100
CTTATAGCCA TCCCCAAAGT TTTGATTATA GTCGTTTCCG ACGGAGACAT 12150
TTTGGGATGG AGCTACACGG TGCTCGGGAA ACGTAACAGT CCGCGCGTAG 12200
TAGTCGAAAC GCACATGCCC TCGAAGGTCC CGATGAACAA AGTAGTAATT 12250
```

FIGURE 1H

GGCAGTCCCG GACCAATGGA CGAAACGGGT AACTATAAAA TGTACTTCGT 12300
CGTCGCGGGG GTGGCCGCGA CGTGCGTAAT TCTTACATGC GCTCTGCTTG 12350
TGGGGAAAAA GAAGTGCCCC GCGCACCAAA TGGGTACTTT TTCCAAGACC 12400
GAACCATTGT ACGCGCCGCT CCCCAAAAAC GAGTTTGAGG CCGGCGGGCT 12450
TACGGACGAT GAGGAAGTGA TTTATGACGA AGTATACGAA CCCCTATTTC 12500
GCGGCTACTG TAAGCAGGAA TTCCGCGAAG ATGTGAATAC CTTTTTCGGT 12550
GCGGTCGTGG AGGGAGAAAG GGCCTTAAAC TTTAAATCCG CCATCGCATC 12600
AATGGCAGAT CGCATCCTGG CAAATAAAAG CGGCAGAAGG AATATGGATA 12650
GCTATTAGTT GGTCATGCCT TTTAAGACCA GAGGGGCCGA AGACGCGGCC 12700
GCGGGCAAGA ACAGGTTTAA GAAATCGAGA AATCGGGAAA TCTTACCGAC 12750
CAGACTGCGT GGCACCGGTA AGAAAACTGC CGGATTGTCC AATTATACCC 12800
AGCCTATTCC CTGGAACCCT AAATTCTGCA GCGCGCGCGG GGAATCTGAC 12850
AACCACGCGT GTAAAGACAC TTTTTATCGC AGGACGTGCT GCGCATCGCG 12900
CTCTACCGTT TCCAGTCAAC CCGATTCCCC CCACACACCC ATGCCTACTG 12950
AGTATGGGCG CGTGCCCTCC GCAAAGCGCA AAAAACTATC ATCTTCAGAC 13000
TSSGAGGGCG CGCACCAACC CCTAGTATCC TGTAAACTTC CGGATTCTCA 13050
AGCAGCACCG GCGCGAACCT ATAGTTCTGC GCAAAGATAT ACTGTTGACG 13100
AGGTTTCGTC GCCAACTCCG CCAGGCGTCG ACGCTGTTGC GGACTTAGAA 13150
ACGCGCGCGG AACTTCCTGG CGCTACGACG GAACAAACGG AAAGTAAAAA 13200
TAAGCTCCCC AACCAACAAT CGCGCCTGAA GCCGAAACCC ACAAACGAGC 13250
ACGTCGGAGG GGAGCGGTGC CCCTCCGAAG GCACGGTCGA GGCGCCATCG 13300
CTCGGCATCC TCTCGCGCGT CGGGGCAGCG ATAGCAAACG AGCTGGCTCG 13350
TATGCGGAGG GCGTGTCTTC CGCTCGCCGC GTCGGCGGCC GCTGCCGGAA 13400
TAGTGGCCTG GGCCGCGGCG AGGGCCTTGC AGAAACAAGG GCGGTAGCAG 13450
TAATAATAAC CACACAAATA TTG 13473

FIGURE 2

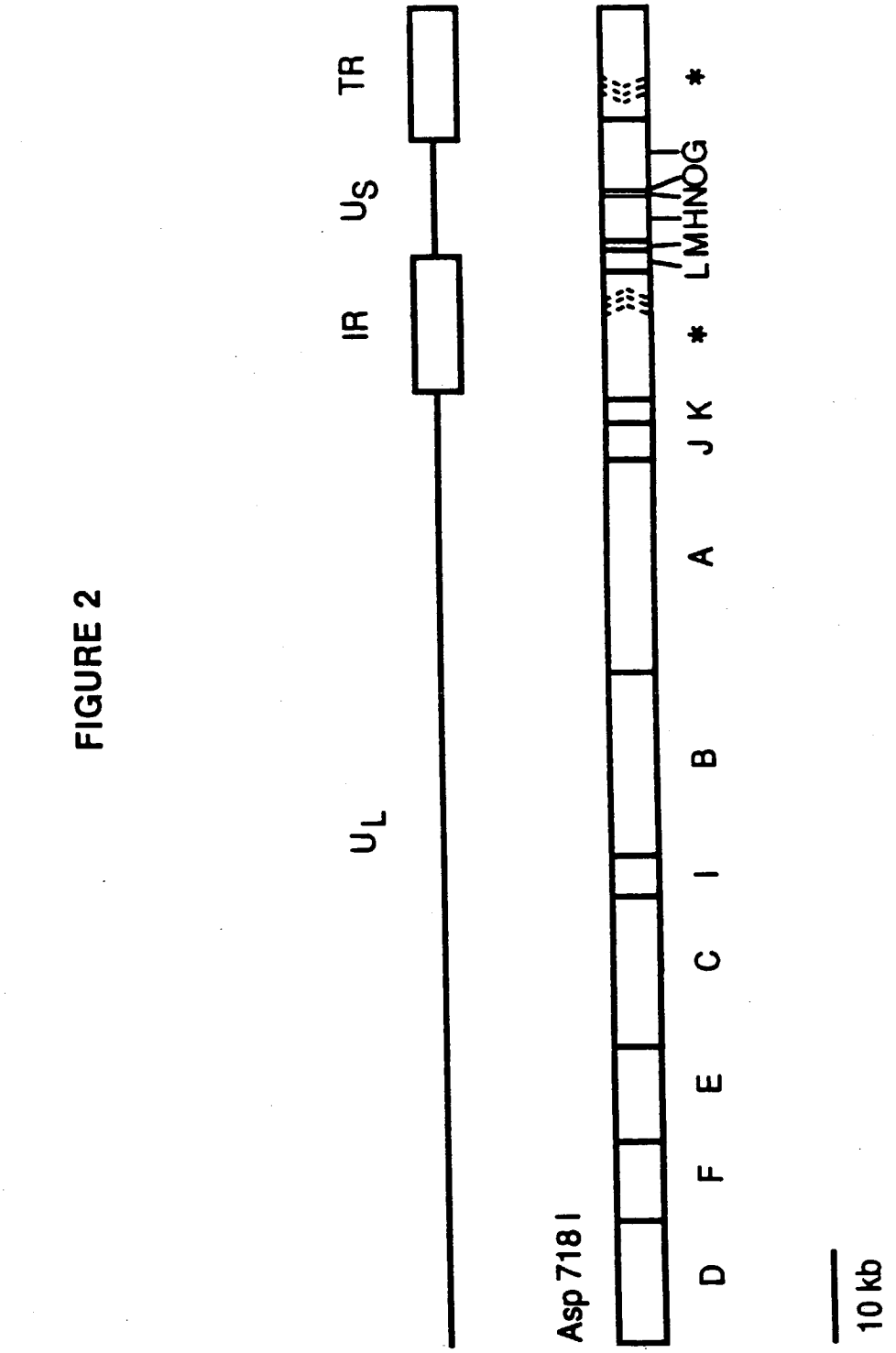


FIGURE 3

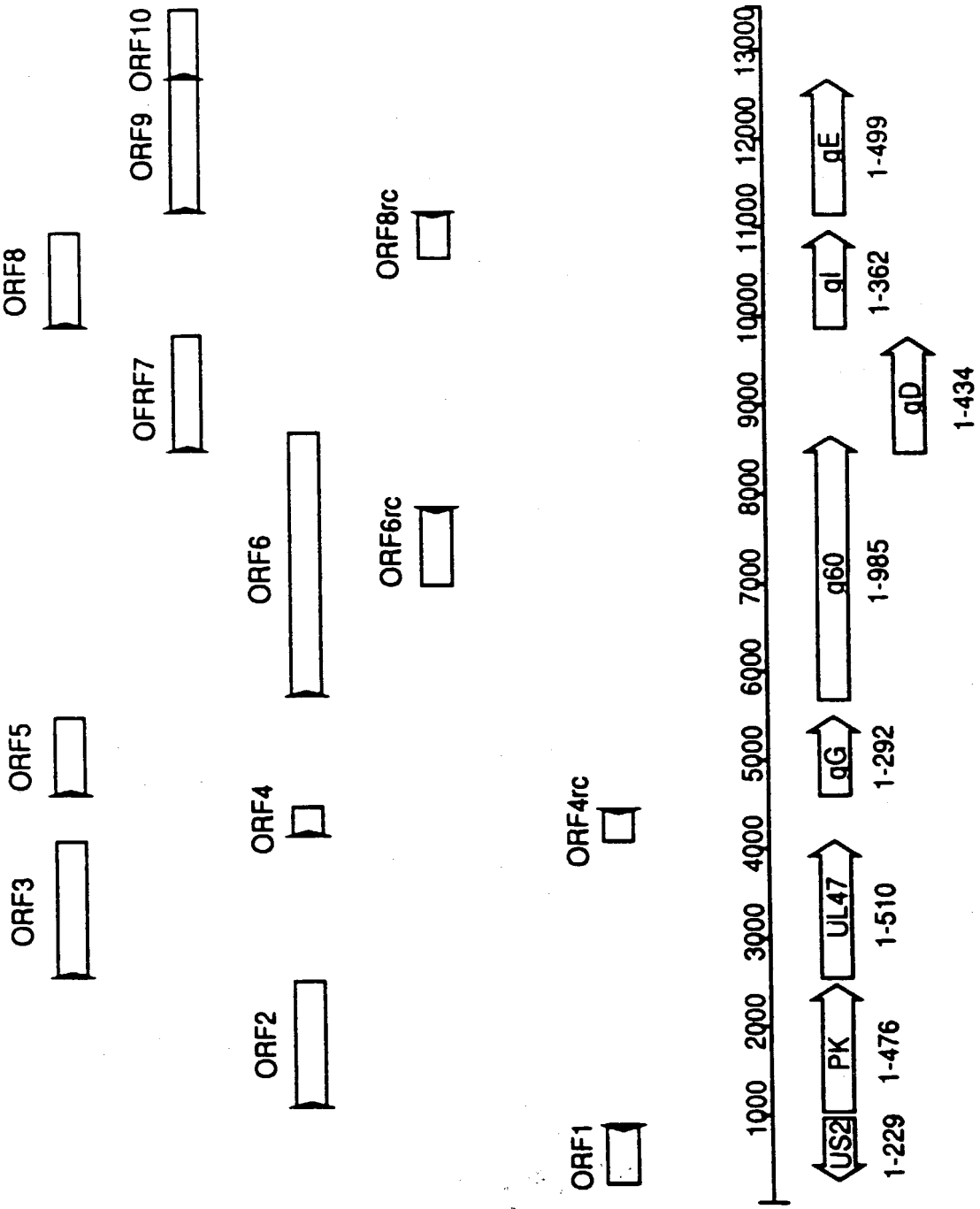

1
2
3
4
5
6
ORF1
ORF2
ORF3
ORF4
ORF4rc
ORF5
ORF6
ORF6rc
OFRF7
ORF8
ORF8rc
ORF9
ORF10
US2
PK
UL47
gG
g60
gD
gI
gE
1-229
1-476
1-510
1-292
1-985
1-434
1-362
1-499
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
11000
12000
13000

FIGURE 4A

FIGURE 4A
FIGURE 4B

DNA	Origin	Sites	Size
Vector	pUC 19	Asp718 I—Asp718 I	~2686 BP
Fragment 1	ILTV 5164 bp Asp718 I	Asp718 I—Nhe I	~2830 BP
Fragment 2	HCMV, E. coli, PRV	Sal I—Sal I	~5017 BP
Fragment 3	ILTV 5164 bp Asp718 I	Sal I—Asp718 I	~1709 BP

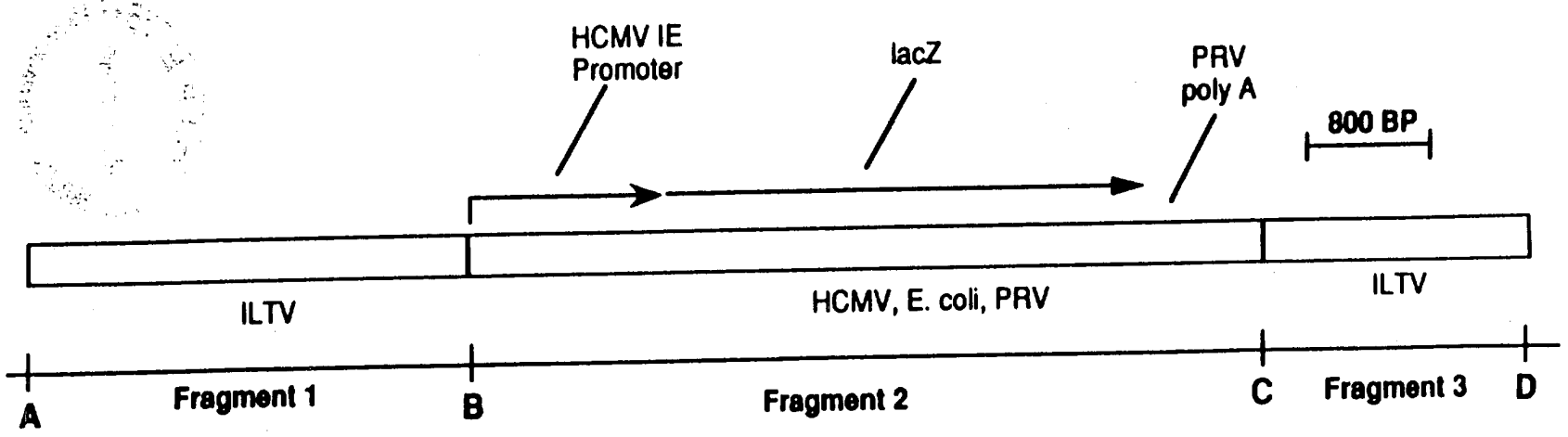


FIGURE 4B

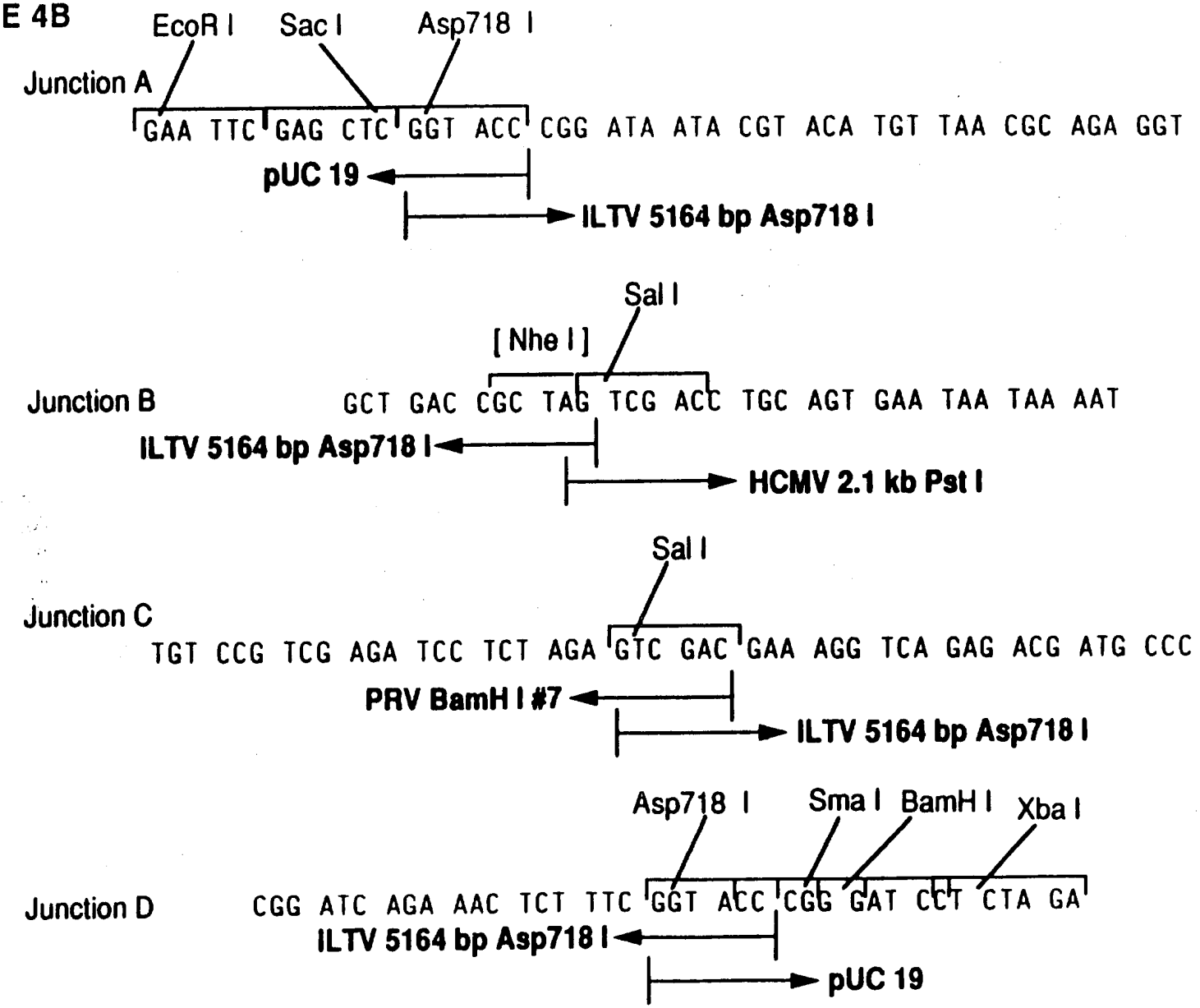

EcoR I
Sac I
Asp718 I
Junction A
GAA TTC GAG CTC GGT ACC CGG ATA ATA CGT ACA TGT TAA CGC AGA GGT
pUC 19
ILTV 5164 bp Asp718 I
Sal I
[Nhe I]
Junction B
GCT GAC CGC TAG TCG ACC TGC AGT GAA TAA TAA AAT
ILTV 5164 bp Asp718 I
HCMV 2.1 kb Pst I
Sal I
Junction C
TGT CCG TCG AGA TCC TCT AGA GTC GAC GAA AGG TCA GAG ACG ATG CCC
PRV BamH I #7
ILTV 5164 bp Asp718 I
Asp718 I
Sma I
BamH I
Xba I
Junction D
CGG ATC AGA AAC TCT TTC GGT ACC CGG GAT CCT CTA GA
ILTV 5164 bp Asp718 I
pUC 19

FIGURE 5A

FIGURE 5A
FIGURE 5B

DNA	Origin	Sites	Size
Vector	pSP 64/65	Hind III—Hind III	~3002 BP
Fragment 1	ILTV 2.4 kb Hind III	Hind III—Bcl I	~1087 BP
Fragment 2	PRV, E. coli, HCMV	Sal I—Sal I	~5017 BP
Fragment 3	ILTV 2.4 kb Hind III	Bcl I—Hind III	~ 700 BP

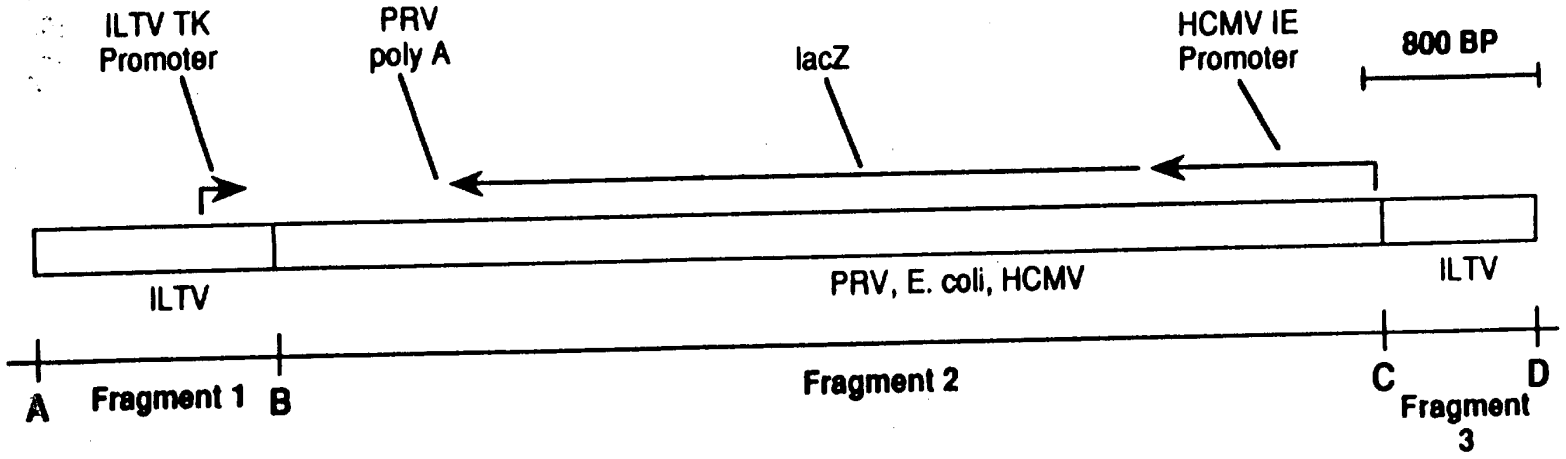


FIGURE 5B

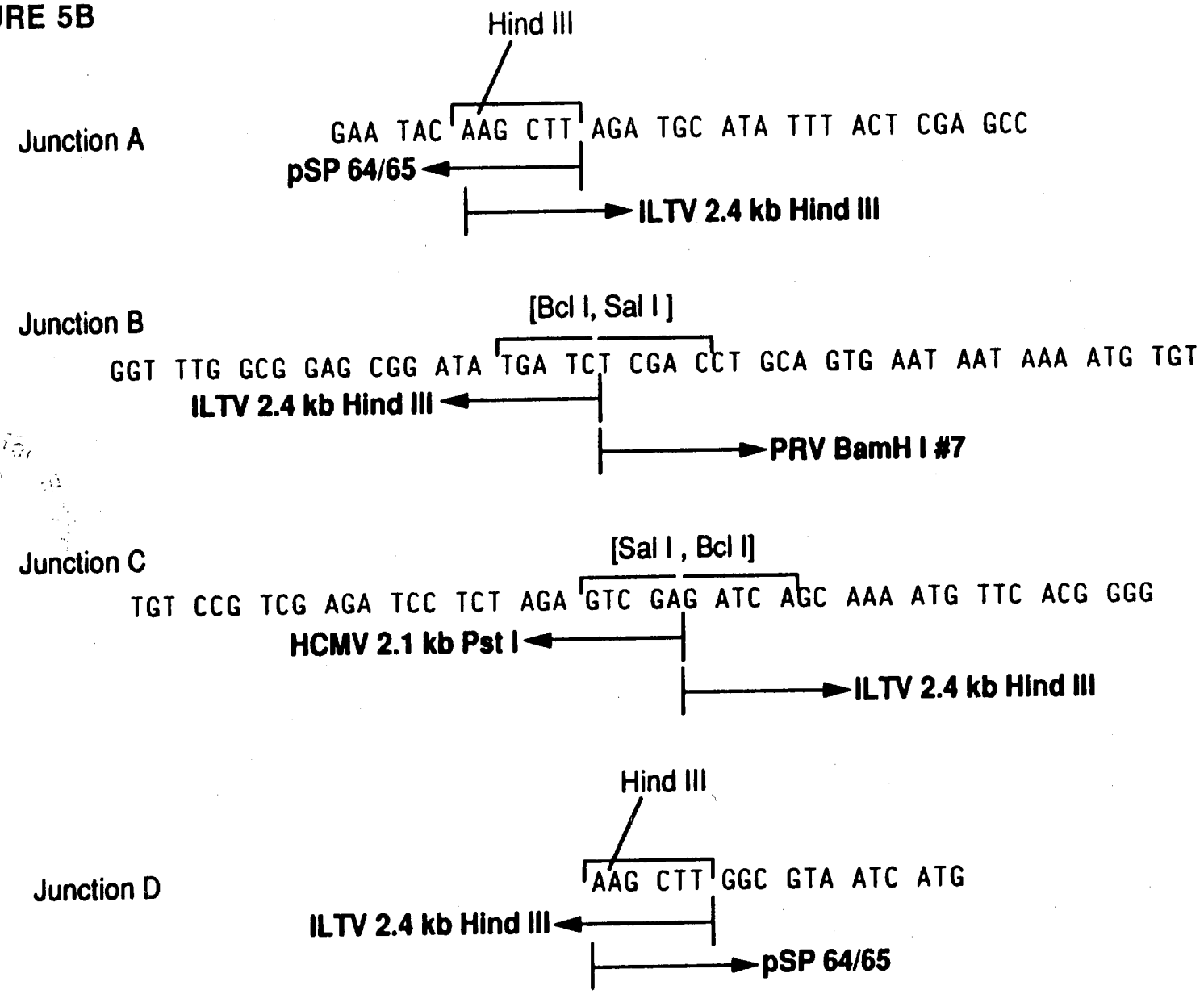

Hind III
Junction A
GAA TAC AAG CTT AGA TGC ATA TTT ACT CGA GCC
pSP 64/65
ILTV 2.4 kb Hind III
[Bcl I, Sal I]
Junction B
GGT TTG GCG GAG CGG ATA TGA TCT CGA CCT GCA GTG AAT AAT AAA ATG TGT
ILTV 2.4 kb Hind III
PRV BamH I #7
[Sal I , Bcl I]
Junction C
TGT CCG TCG AGA TCC TCT AGA GTC GAG ATC AGC AAA ATG TTC ACG GGG
HCMV 2.1 kb Pst I
ILTV 2.4 kb Hind III
Hind III
Junction D
AAG CTT GGC GTA ATC ATG
ILTV 2.4 kb Hind III
pSP 64/65

FIGURE 6A

FIGURE 6A
FIGURE 6B

DNA	Origin	Sites	Size
Vector	pSP 18/19	Asp718 I—Asp718 I	~2958 BP
Fragment 1	ILTV 2.5 kb Asp718 I	Asp718 I—Dra I	~2300 BP
Fragment 2	PRV, E. coli, HSV-1	Xba I—Xba I	~3039 BP
Fragment 3	ILTV 1097 bp Asp718 I	Xba I—Asp718 I	~ 809 BP

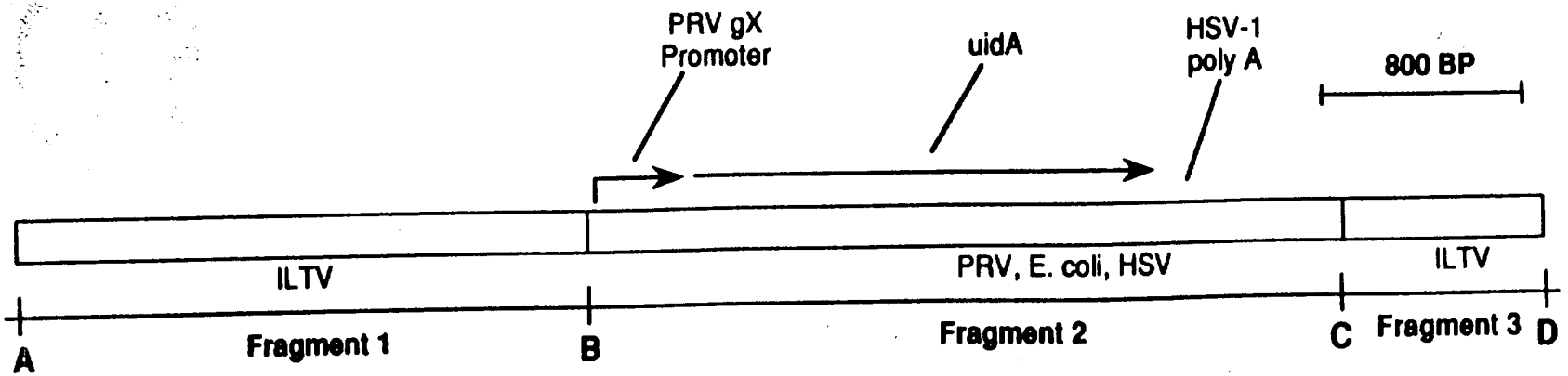


FIGURE 6B

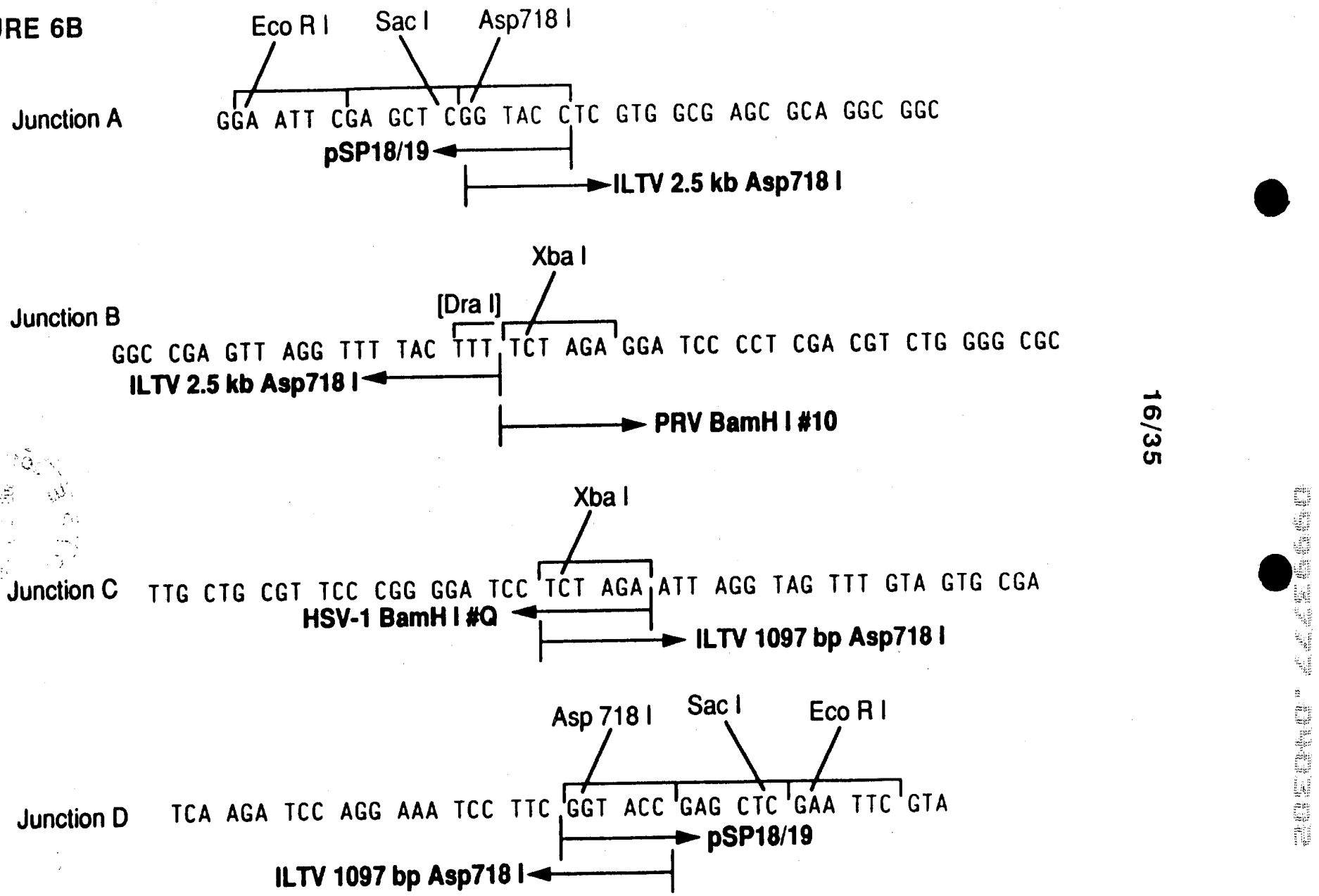

Eco R I
Sac I
Asp718 I
Junction A
GGA ATT CGA GCT CGG TAC CTC GTG GCG AGC GCA GGC GGC
pSP18/19
ILTV 2.5 kb Asp718 I
Xba I
[Dra I]
Junction B
GGC CGA GTT AGG TTT TAC TTT TCT AGA GGA TCC CCT CGA CGT CTG GGG CGC
ILTV 2.5 kb Asp718 I
PRV BamH I #10
Xba I
Junction C
TTG CTG CGT TCC CGG GGA TCC TCT AGA ATT AGG TAG TTT GTA GTG CGA
HSV-1 BamH I #Q
ILTV 1097 bp Asp718 I
Asp 718 I
Sac I
Eco R I
Junction D
TCA AGA TCC AGG AAA TCC TTC GGT ACC GAG CTC GAA TTC GTA
pSP18/19
ILTV 1097 bp Asp718 I

FIGURE 7A

DNA	Origin	Sites	Size
Vector	pUC19	Asp718 I—Hind III	~2647 BP
Fragment 1	ILTV 8.0 kb Asp718 I	Asp718 I—Xba I	~1619 BP
Fragment 2	ILTV 8.0 kb Asp718 I	Xba I—Xho I†	~ 691 BP
Fragment 3	HSV-1, E. coli, PRV	Sal I—Sal I	~3051 BP
Fragment 4	ILTV 8.0 kb Asp718 I	Xho I†—Hind III	~ 624 BP
Fragment 5	ILTV 8.0 kb Asp718 I	Hind III—Hind III	~2700 BP

†Restriction enzyme site introduced by PCR cloning

FIGURE 7A
FIGURE 7B
FIGURE 7C

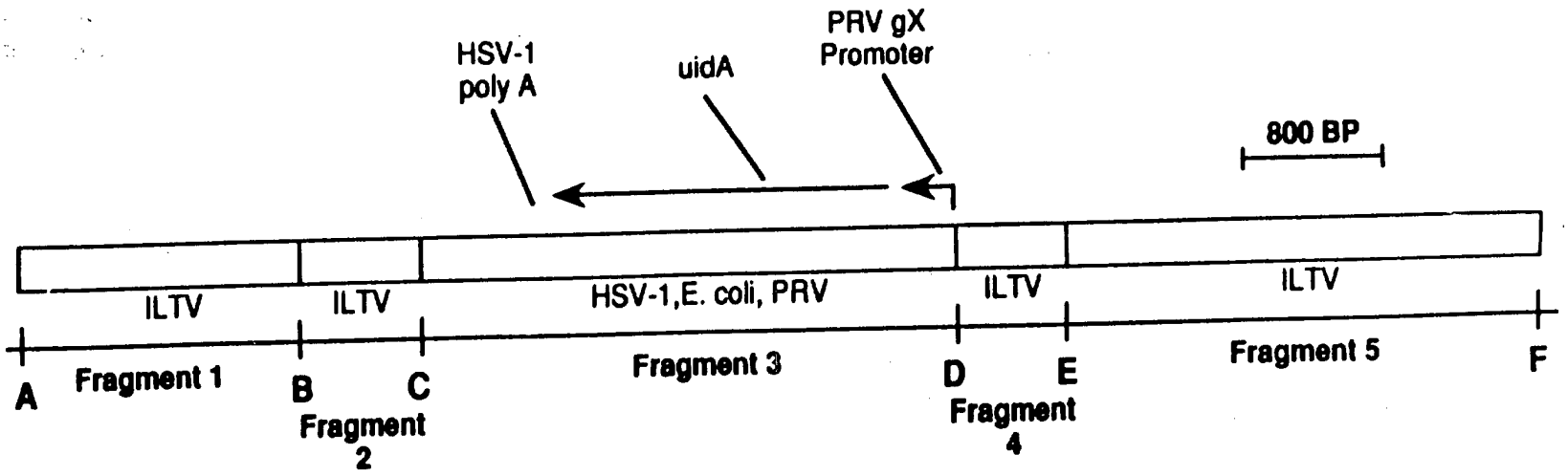


FIGURE 7B

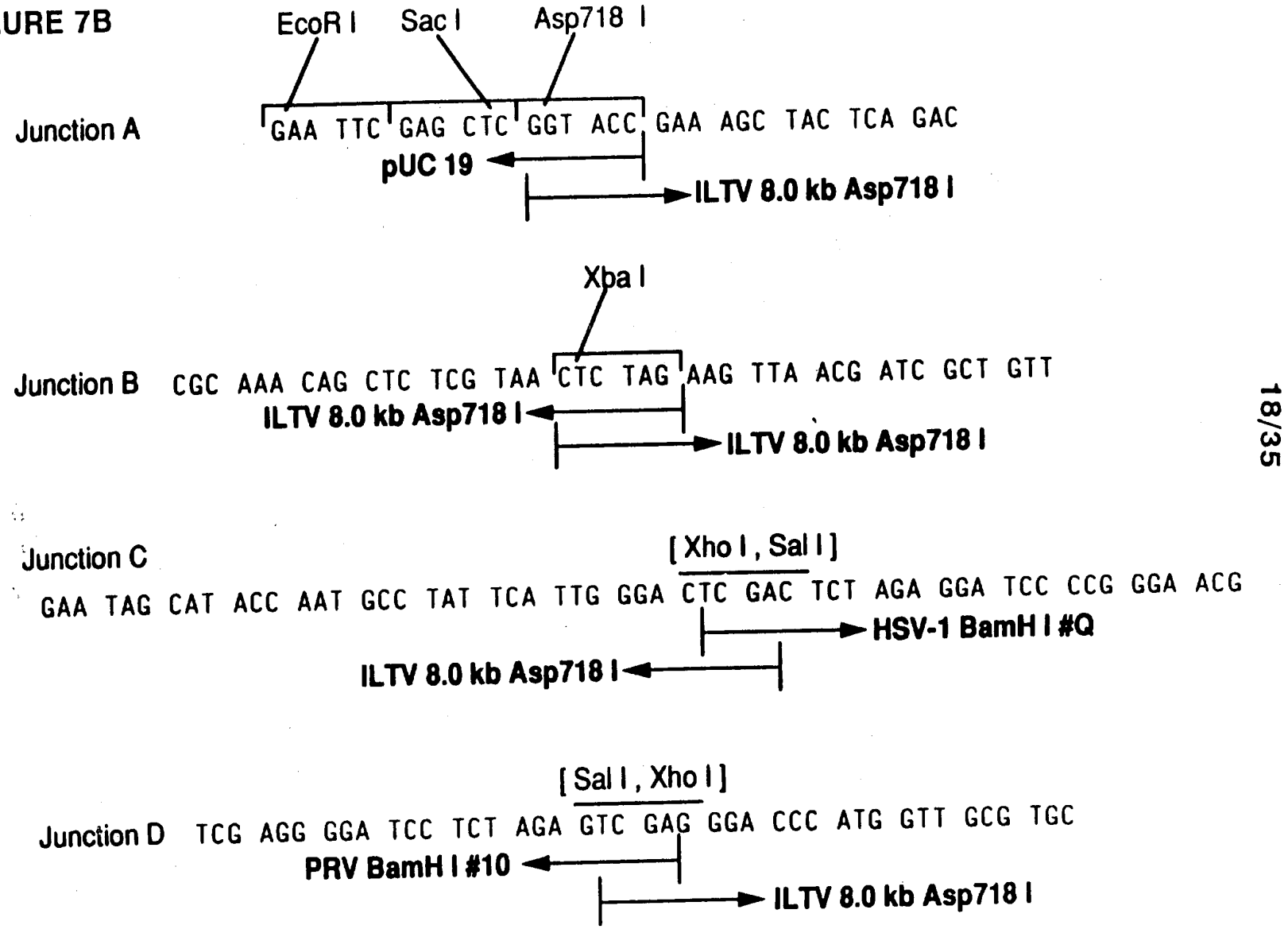

EcoR I
Sac I
Asp718 I
Junction A
GAA TTC GAG CTC GGT ACC GAA AGC TAC TCA GAC
pUC 19
ILTV 8.0 kb Asp718 I
Xba I
Junction B
CGC AAA CAG CTC TCG TAA CTC TAG AAG TTA ACG ATC GCT GTT
ILTV 8.0 kb Asp718 I
ILTV 8.0 kb Asp718 I
Junction C
[Xho I, Sal I]
GAA TAG CAT ACC AAT GCC TAT TCA TTG GGA CTC GAC TCT AGA GGA TCC CCG GGA ACG
HSV-1 BamH I #Q
ILTV 8.0 kb Asp718 I
[Sal I, Xho I]
Junction D
TCG AGG GGA TCC TCT AGA GTC GAG GGA CCC ATG GTT GCG TGC
PRV BamH I #10
ILTV 8.0 kb Asp718 I

FIGURE 7C

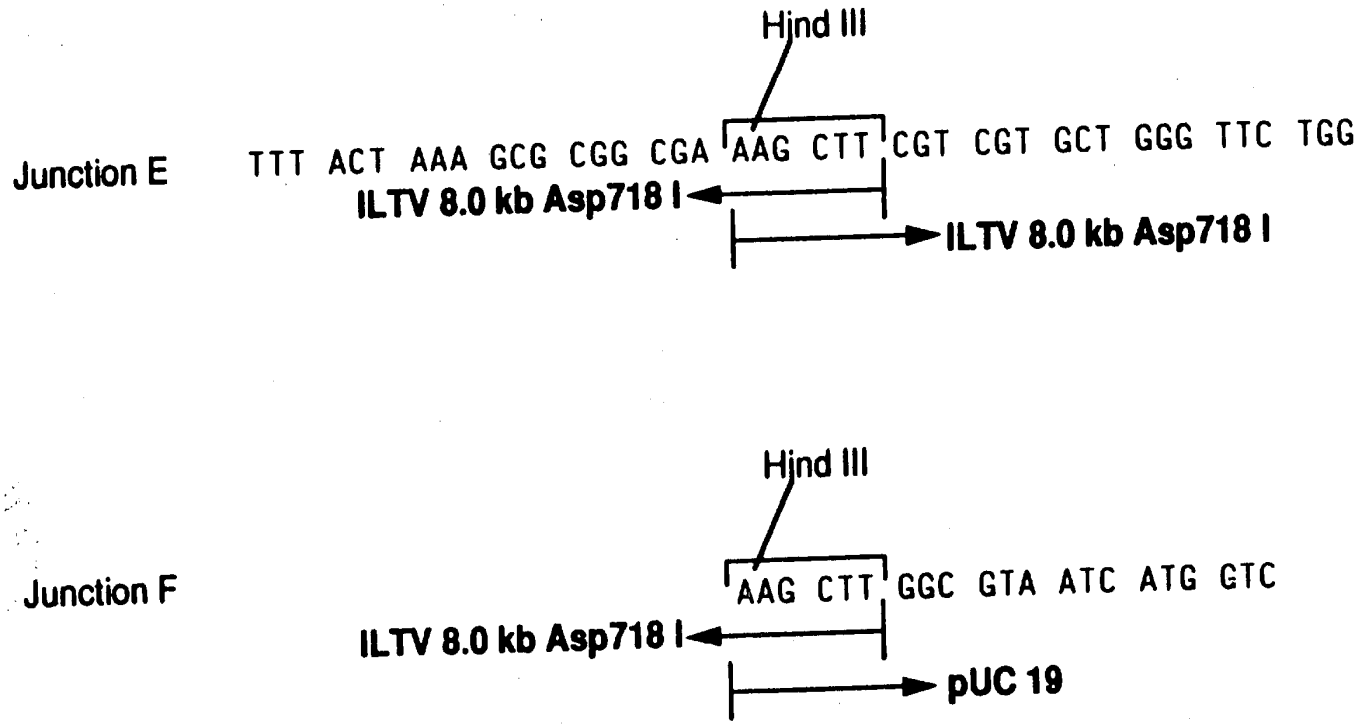

Hind III
Junction E
TTT ACT AAA GCG CGG CGA AAG CTT CGT CGT GCT GGG TTC TGG
ILTV 8.0 kb Asp718 I
ILTV 8.0 kb Asp718 I
Hind III
Junction F
AAG CTT GGC GTA ATC ATG GTC
ILTV 8.0 kb Asp718 I
pUC 19

FIGURE 8A
FIGURE 8B
FIGURE 8C

FIGURE 8A

DNA	Origin	Sites	Size
Vector	pSP18/19	Asp718 I—Asp718 I	~2958 BP
Fragment 1	ILTV 5164 bp Asp718I	Asp718 I—BssH II	~1066 BP
Fragment 2	ILTV 5164 bp Asp718I	Sal I—Bcl I	~ 123 BP
Fragment 3	HSV-1, E. coli, PRV	BamH I—BamH I	~3027 BP
Fragment 4	ILTV 5164 bp Asp718I	Bcl I—Asp718 I	~1334 BP

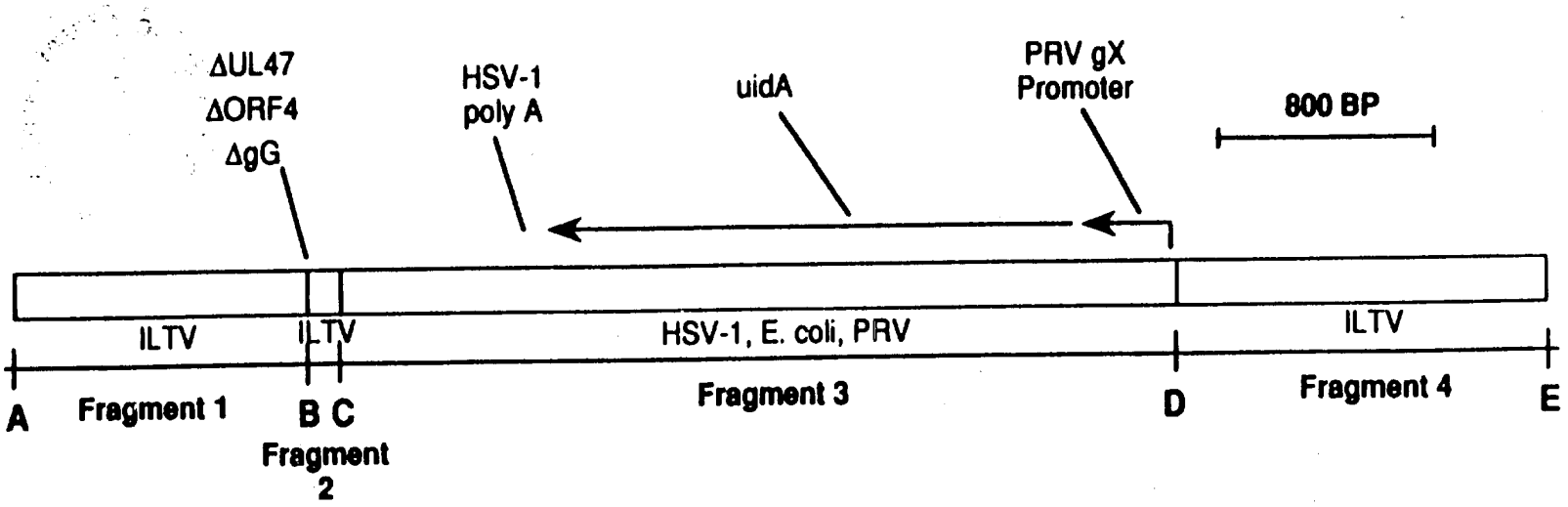


FIGURE 8B

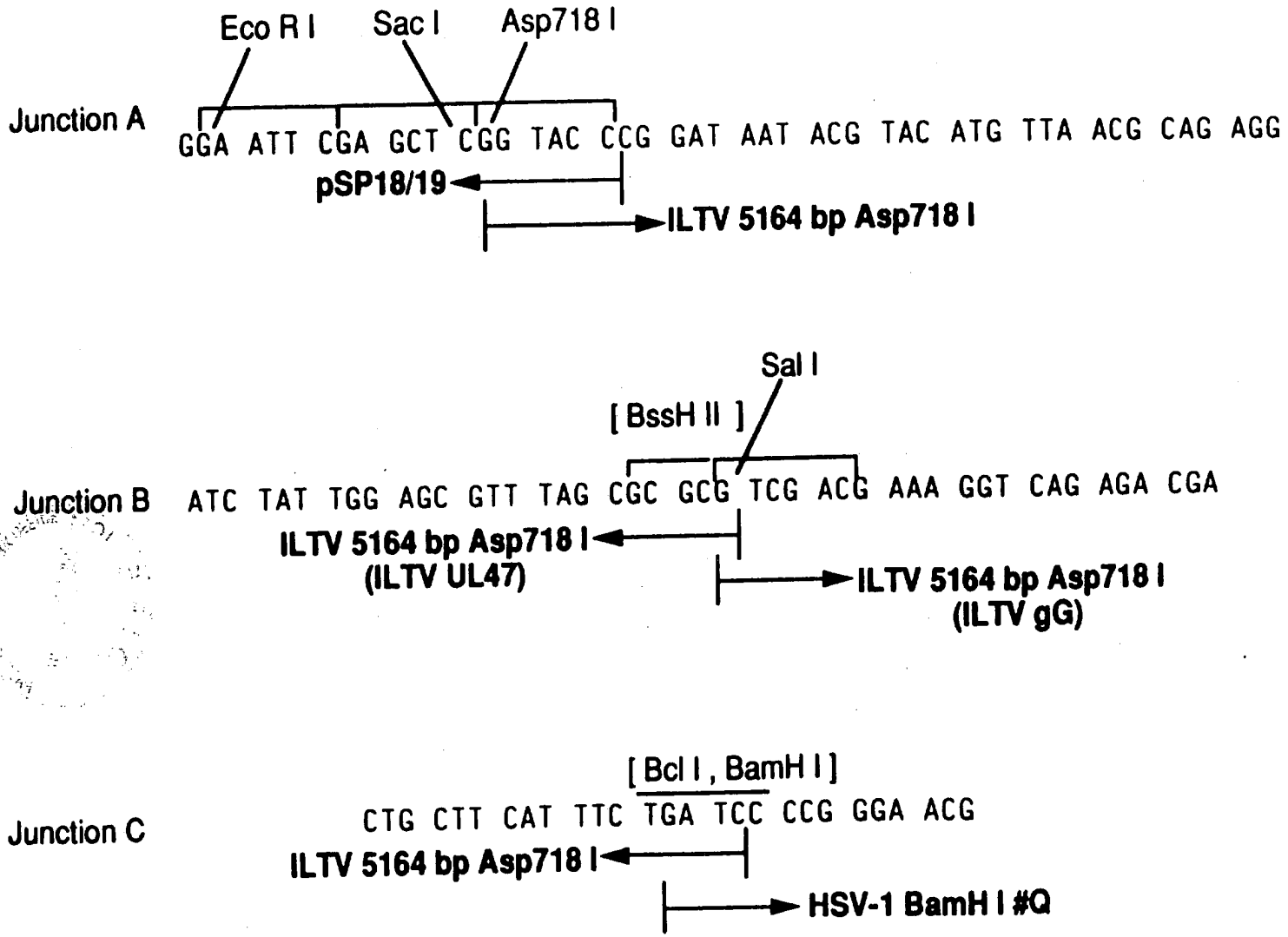

Eco R I
Sac I
Asp718 I
Junction A
GGA ATT CGA GCT CGG TAC CCG GAT AAT ACG TAC ATG TTA ACG CAG AGG
pSP18/19
ILTV 5164 bp Asp718 I
Sal I
[BssH II]
Junction B
ATC TAT TGG AGC GTT TAG CGC GCG TCG ACG AAA GGT CAG AGA CGA
ILTV 5164 bp Asp718 I
(ILTV UL47)
ILTV 5164 bp Asp718 I
(ILTV gG)
[Bcl I , BamH I]
Junction C
CTG CTT CAT TTC TGA TCC CCG GGA ACG
ILTV 5164 bp Asp718 I
HSV-1 BamH I #Q

FIGURE 8C

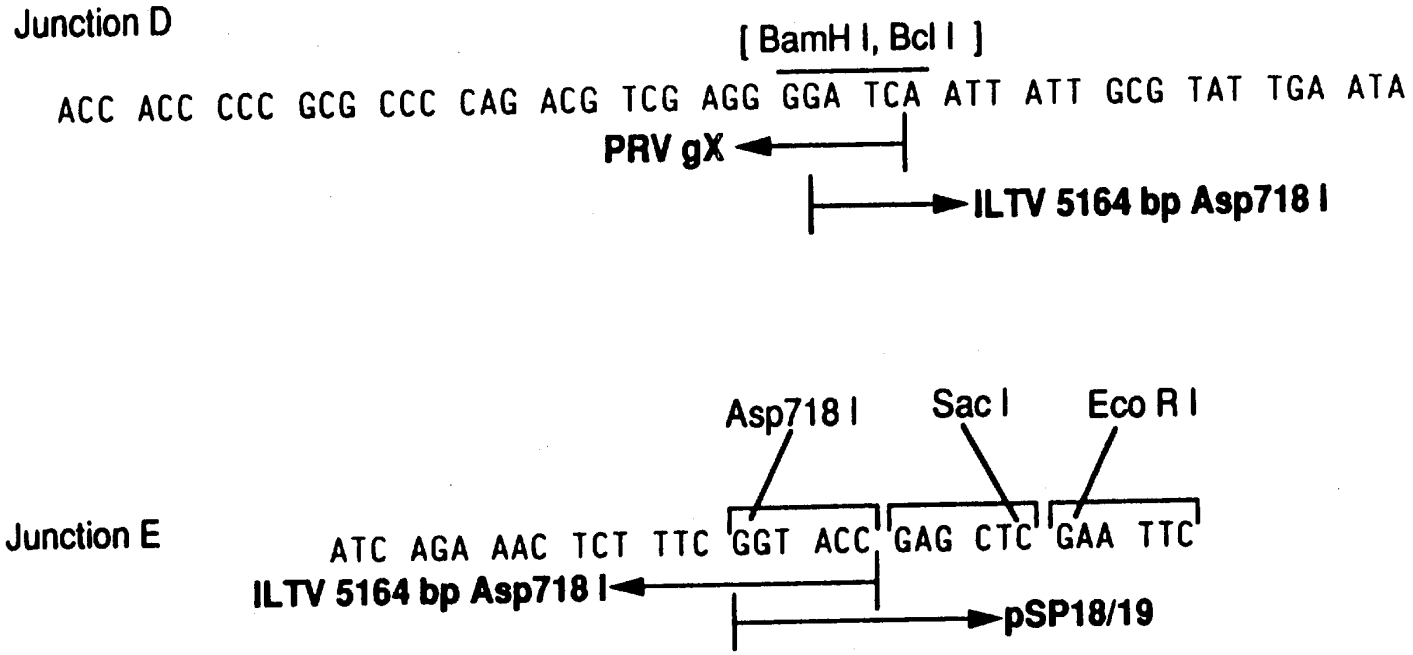

Junction D
[BamH I, Bcl I]
ACC ACC CCC GCG CCC CAG ACG TCG AGG GGA TCA ATT ATT GCG TAT TGA ATA
PRV gX
ILTV 5164 bp Asp718 I
Asp718 I
Sac I
Eco R I
Junction E
ATC AGA AAC TCT TTC GGT ACC GAG CTC GAA TTC
ILTV 5164 bp Asp718 I
pSP18/19

FIGURE 9A

FIGURE 9A
FIGURE 9B

DNA	Origin	Sites	Size
Vector	pUC19	Asp718 I—BamH I	~2677 BP
Fragment 1	ILTV 5164 bp Asp718I	Asp718 I—Nhe I	~2830 BP
Fragment 2	PRV, E. coli, HSV-1	Sal I—Sal I	~3051 BP
Fragment 3	ILTV 4545 bp BamH I	Sal I—BamH I	~1709 BP

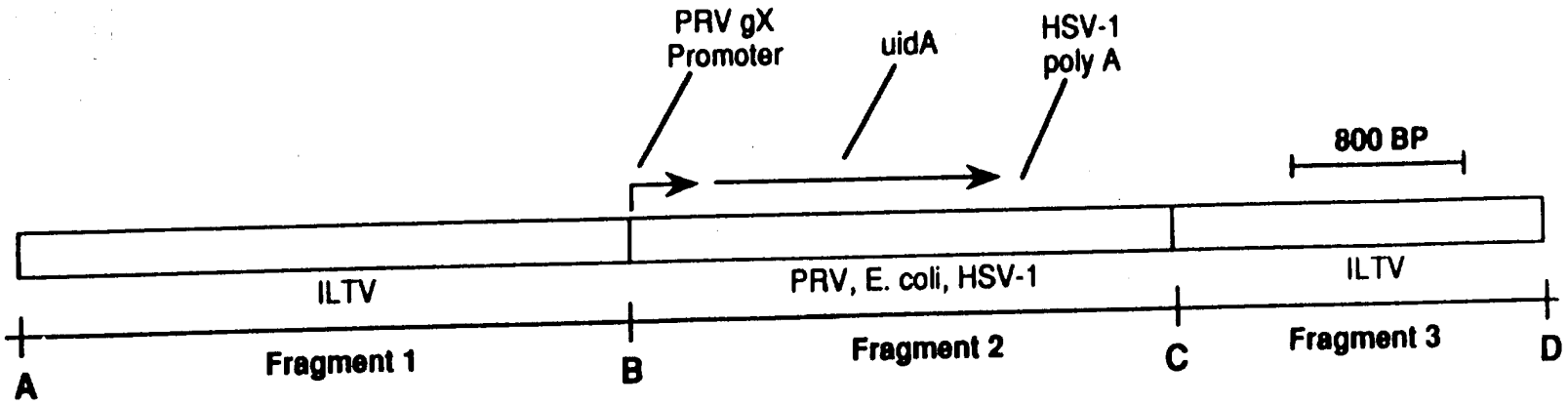


FIGURE 9B

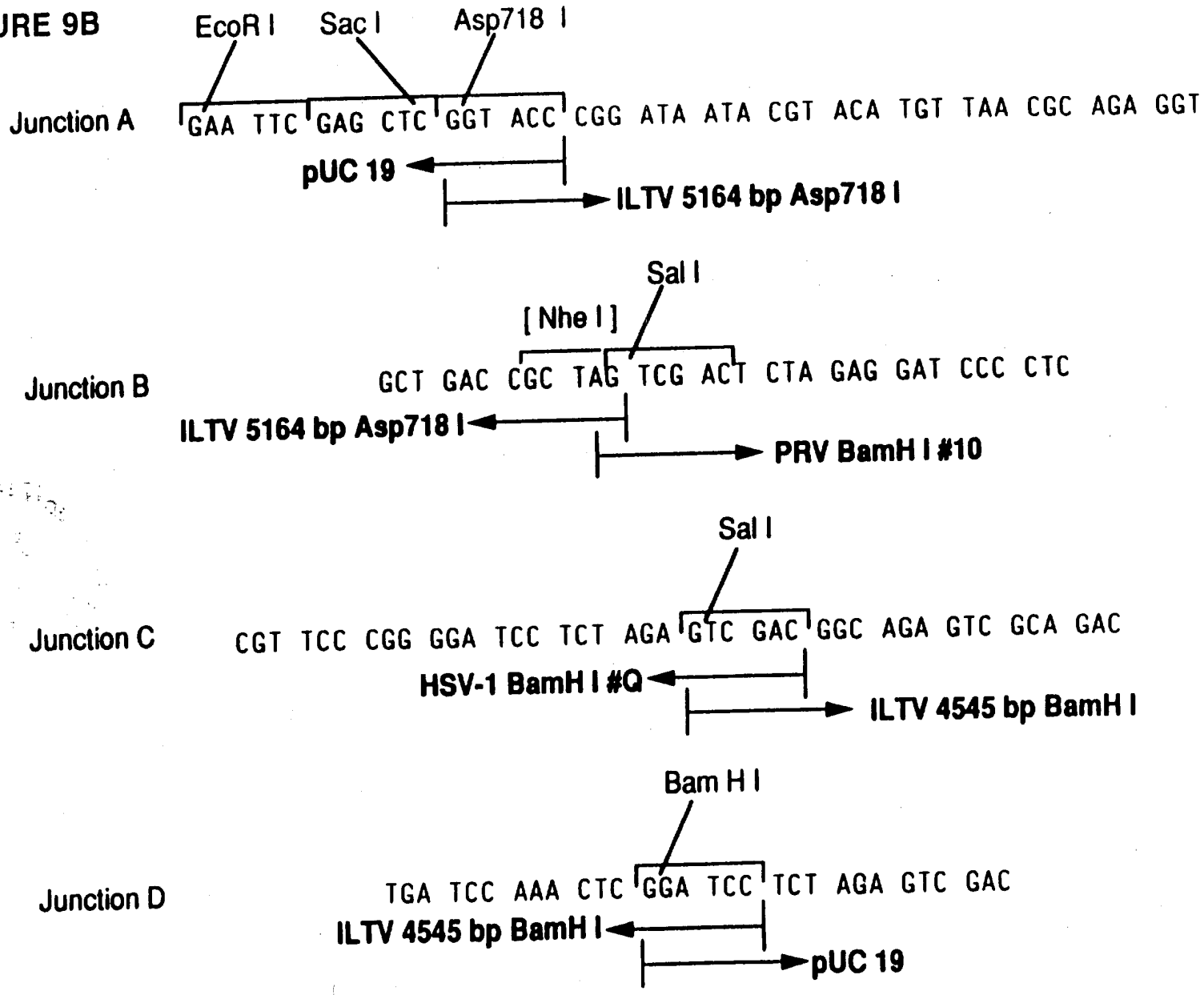

EcoR I
Sac I
Asp718 I
Junction A
GAA TTC GAG CTC GGT ACC CGG ATA ATA CGT ACA TGT TAA CGC AGA GGT
pUC 19
ILTV 5164 bp Asp718 I
Sal I
[Nhe I]
Junction B
GCT GAC CGC TAG TCG ACT CTA GAG GAT CCC CTC
ILTV 5164 bp Asp718 I
PRV BamH I #10
Sal I
Junction C
CGT TCC CGG GGA TCC TCT AGA GTC GAC GGC AGA GTC GCA GAC
HSV-1 BamH I #Q
ILTV 4545 bp BamH I
Bam H I
Junction D
TGA TCC AAA CTC GGA TCC TCT AGA GTC GAC
ILTV 4545 bp BamH I
pUC 19

FIGURE 10A
FIGURE 10B

FIGURE 10A

DNA	Origin	Sites	Size
Vector	pSP 71	Xma I—Sma I	~3066 BP
Fragment 1	PRV BamH I #10	Sal I—EcoR I†	~ 422 BP
Fragment 2	pRAJ 260	EcoR I†—Xma I†	~1826 BP
Fragment 3	HSV-1 BamH I #Q	Xma I—Xma I	~ 784 BP

†Restriction enzyme site introduced by PCR cloning

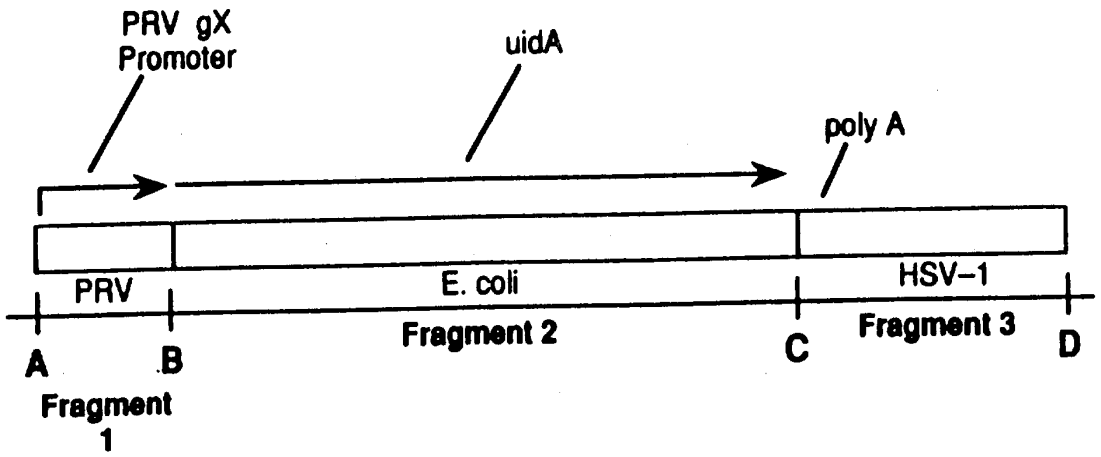


FIGURE 10B

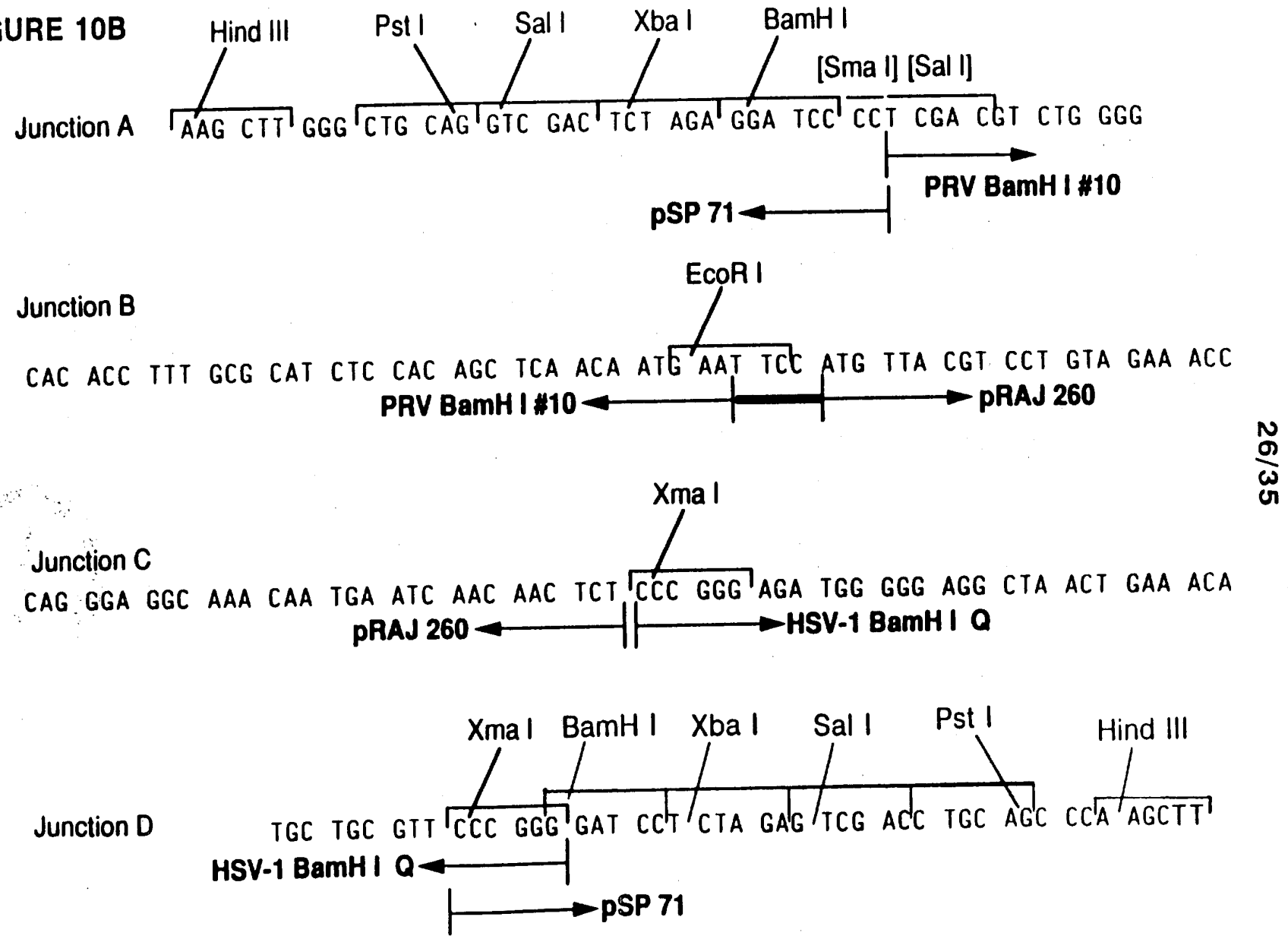

Hind III
Pst I
Sal I
Xba I
BamH I
[Sma I] [Sal I]
Junction A
AAG CTT GGG CTG CAG GTC GAC TCT AGA GGA TCC CCT CGA CGT CTG GGG
PRV BamH I #10
pSP 71
EcoR I
Junction B
CAC ACC TTT GCG CAT CTC CAC AGC TCA ACA ATG AAT TCC ATG TTA CGT CCT GTA GAA ACC
PRV BamH I #10
pRAJ 260
Xma I
Junction C
CAG GGA GGC AAA CAA TGA ATC AAC AAC TCT CCC GGG AGA TGG GGG AGG CTA ACT GAA ACA
pRAJ 260
HSV-1 BamH I Q
Xma I
BamH I
Xba I
Sal I
Pst I
Hind III
Junction D
TGC TGC GTT CCC GGG GAT CCT CTA GAG TCG ACC TGC AGC CCA AGCTT
HSV-1 BamH I Q
pSP 71

FIGURE 11A

DNA	Origin	Sites	Size
Vector	pSP 72	Pst I—Pst I	~3076 BP
Fragment 1	HCMV 2.1 kb Pst I	Pst I—Ava II	~1154 BP
Fragment 2	pJF 751	BamH I—Pvu II	~3010 BP
Fragment 3	PRV BamH I #7	Nde I—Sal I	~ 750 BP

FIGURE 11A
FIGURE 11B
FIGURE 11C

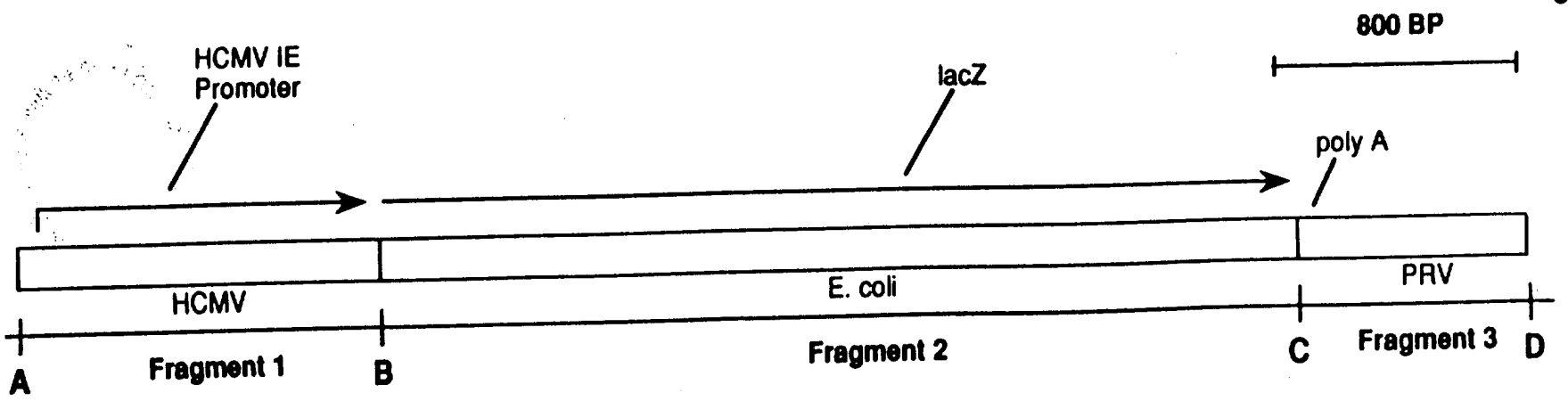


FIGURE 11B

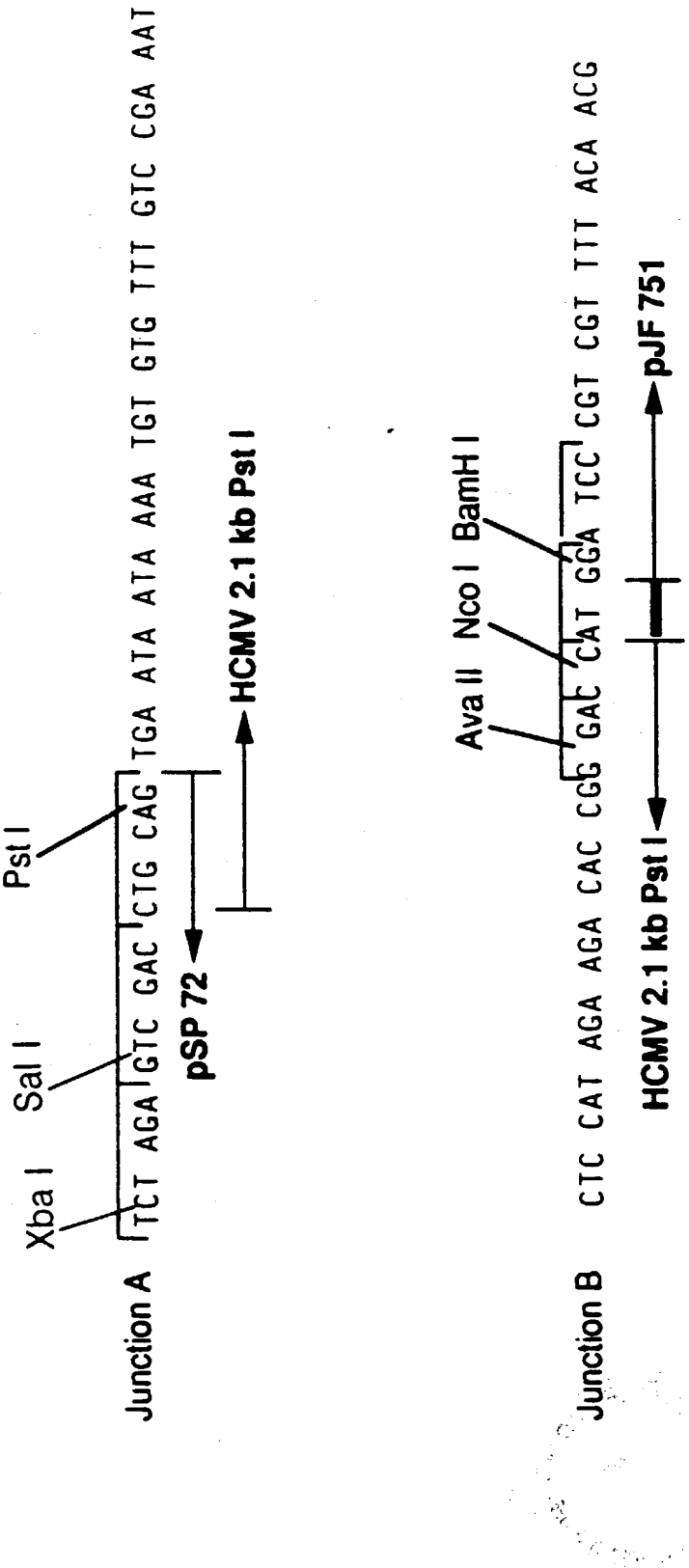
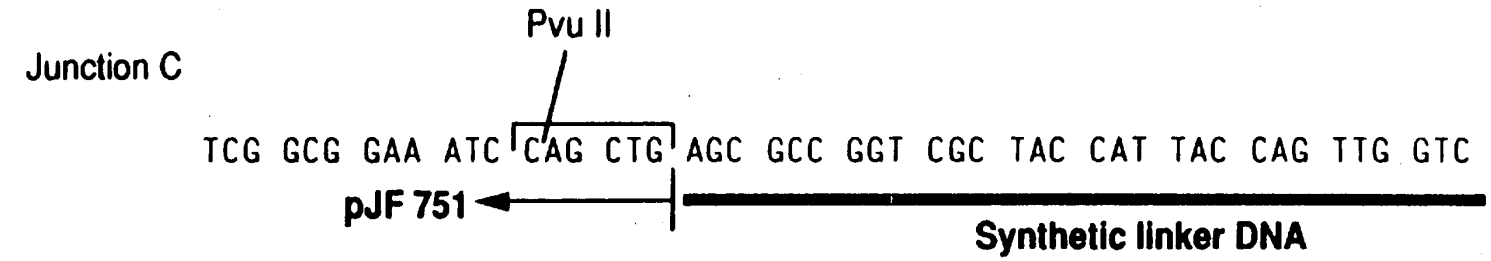
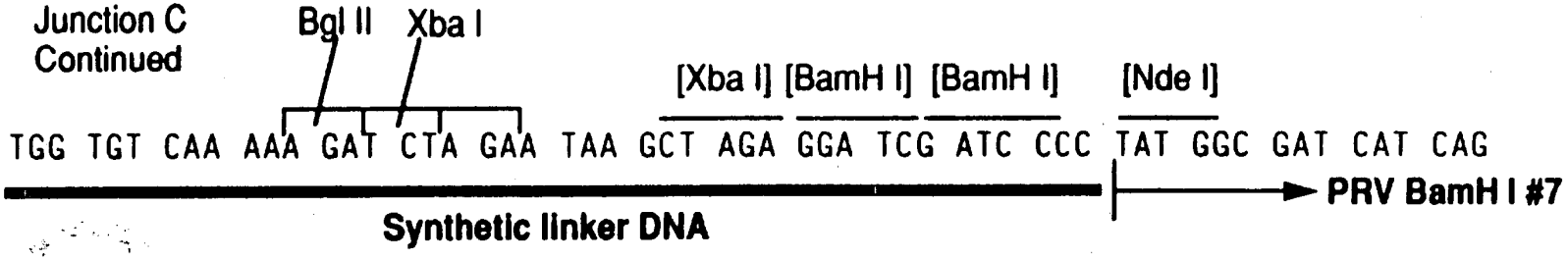

Junction A
Xba I
Sal I
Pst I
TCT AGA GTC GAC CTG CAG TGA ATA ATA AAA TGT GTG TTT GTC CGA AAT
pSP 72
HCMV 2.1 kb Pst I
Junction B
CTC CAT AGA AGA CAC CGG GAC CAT GGA TCC CGT CGT TTT ACA ACG
Ava II
Nco I
BamH I
HCMV 2.1 kb Pst I
pJF 751

FIGURE 11C


Junction C
Pvu II
TCG GCG GAA ATC CAG CTG AGC GCC GGT CGC TAC CAT TAC CAG TTG GTC
pJF 751
Synthetic linker DNA


Junction C
Continued
Bgl II
Xba I
[Xba I] [BamH I] [BamH I]
[Nde I]
TGG TGT CAA AAA GAT CTA GAA TAA GCT AGA GGA TCG ATC CCC TAT GGC GAT CAT CAG
Synthetic linker DNA
PRV BamH I #7

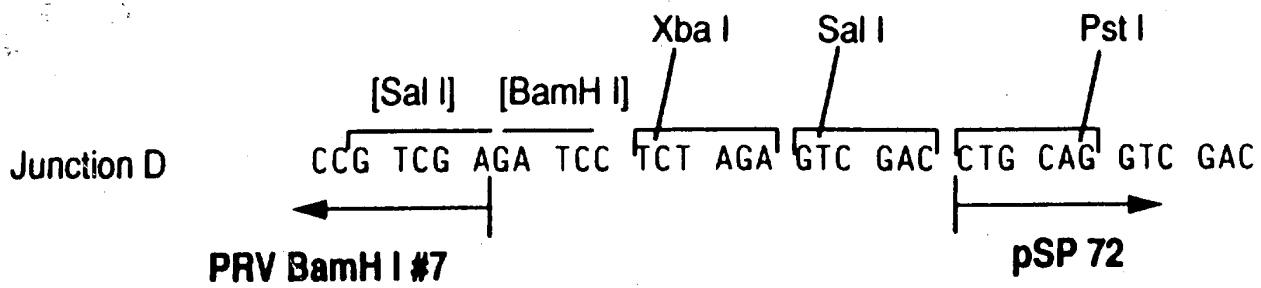

Xba I
Sal I
Pst I
[Sal I]
[BamH I]
Junction D
CCG TCG AGA TCC TCT AGA GTC GAC CTG CAG GTC GAC
PRV BamH I #7
pSP 72

FIGURE 12

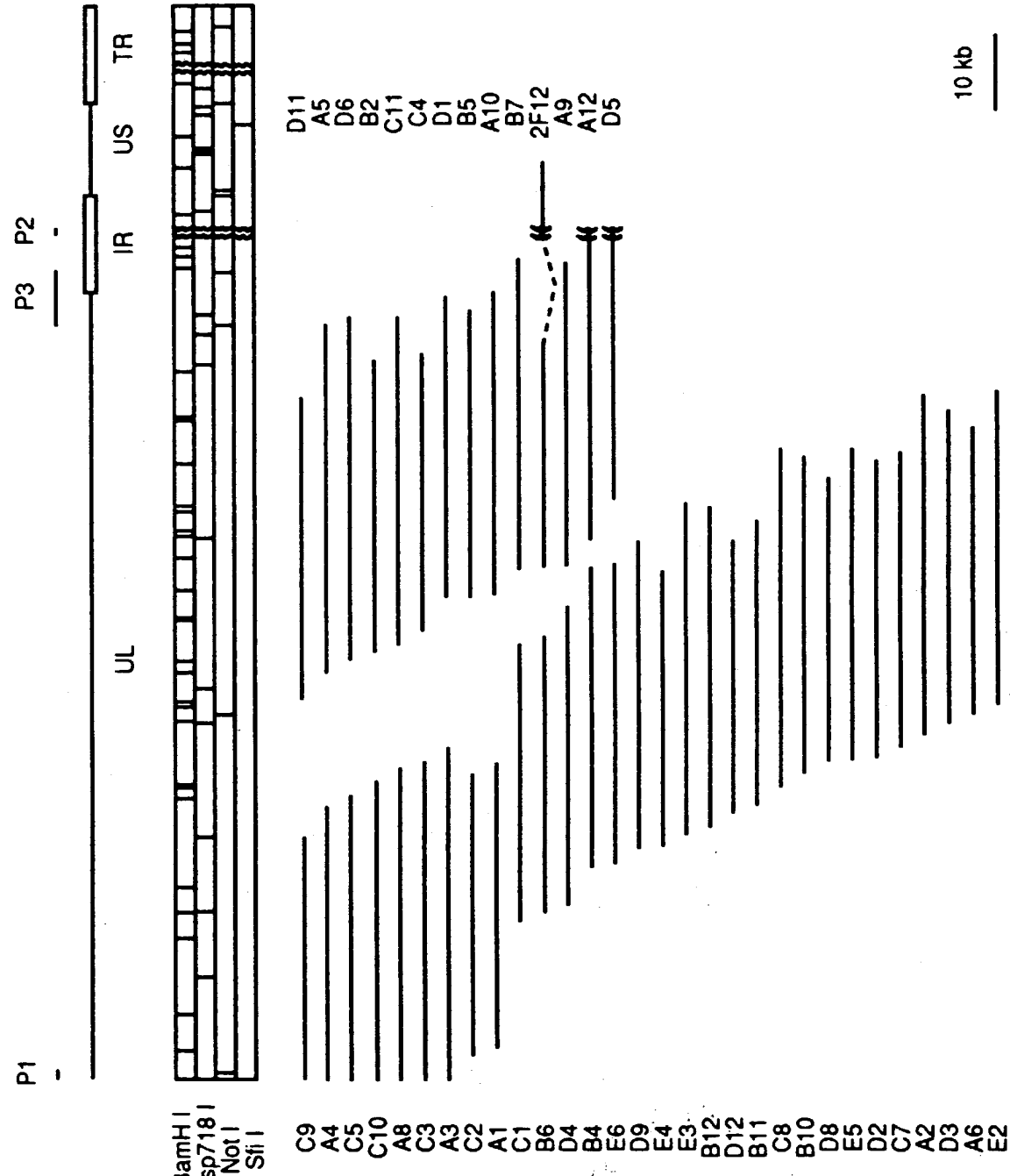

P1
P3
P2
UL
IR
US
TR
BamH I
Asp718 I
Not I
Sfi I
C9
A4
C5
C10
A8
C3
A3
C2
A1
C1
B6
D4
B4
E6
D9
E4
E3
B12
D12
B11
C8
B10
D8
E5
D2
C7
A2
D3
A6
E2
D11
A5
D6
B2
C11
C4
D1
B5
A10
B7
2F12
A9
A12
D5
10 kb

FIGURE 13

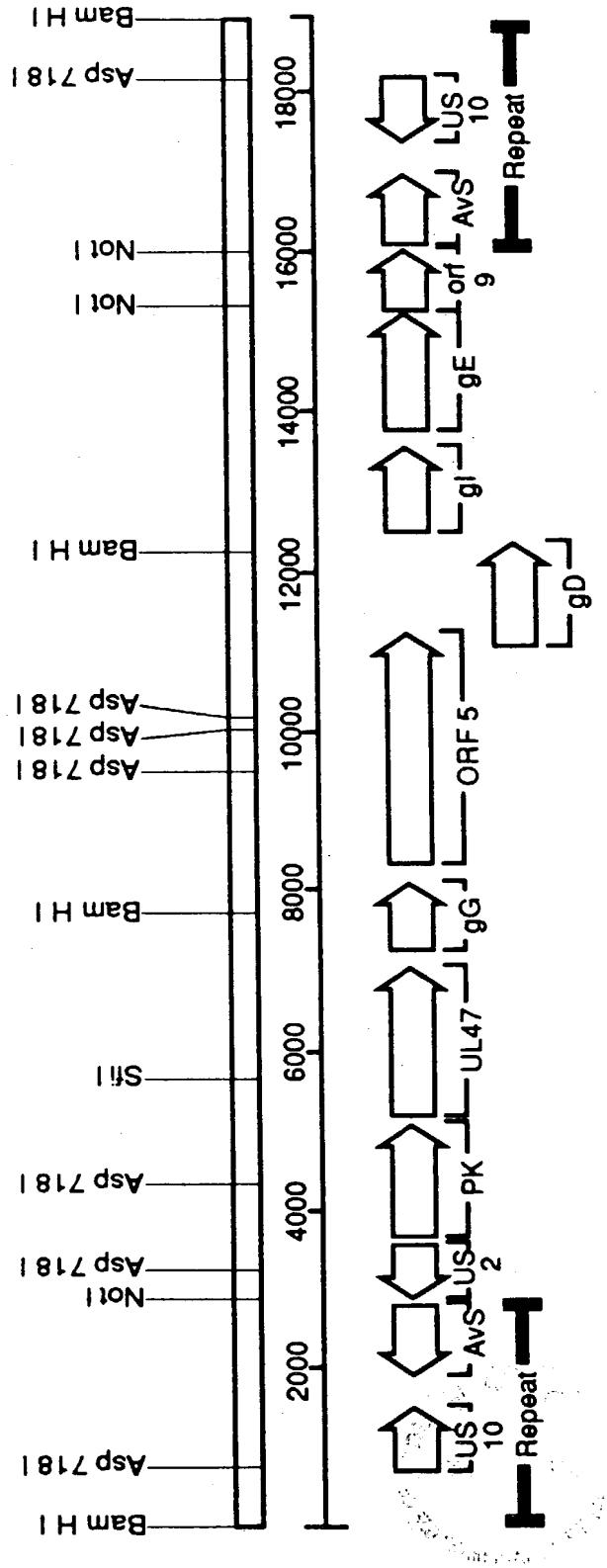

Bam H I
Asp 718 I
Not I
Not I
Bam H I
Asp 718 I
Asp 718 I
Asp 718 I
Bam H I
Sfi I
Asp 718 I
Asp 718 I
Not I
Asp 718 I
Bam H I
2000
4000
6000
8000
10000
12000
14000
16000
18000
US 10
AvS
orf 9
gE
gI
gD
ORF 5
gG
UL47
PK
US 2
AvS
US 10
Repeat
Repeat

FIGURE 14

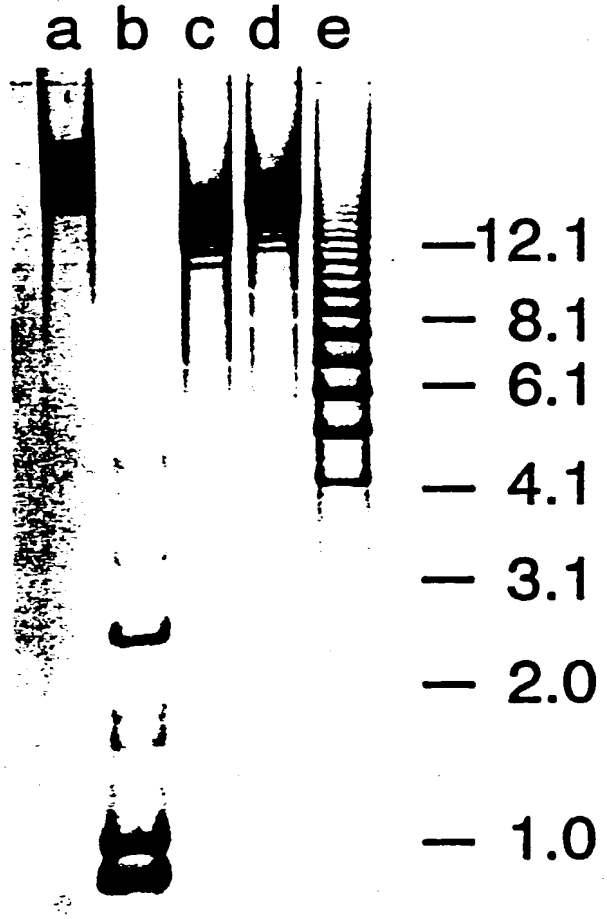


FIGURE 15

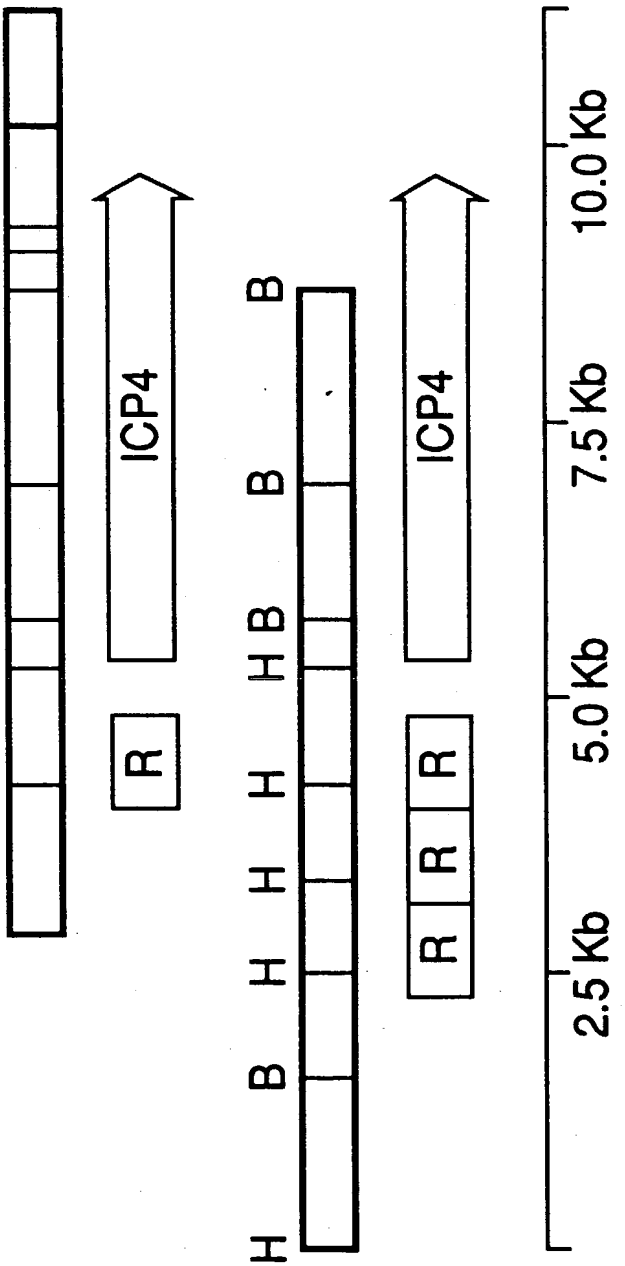

SA-2
B
H
H B
B
B H N
H
R
ICP4
USDA
H
B
B H H H
H B
B
R R R
ICP4
2.5 Kb
5.0 Kb
7.5 Kb
10.0 Kb

FIGURE 16

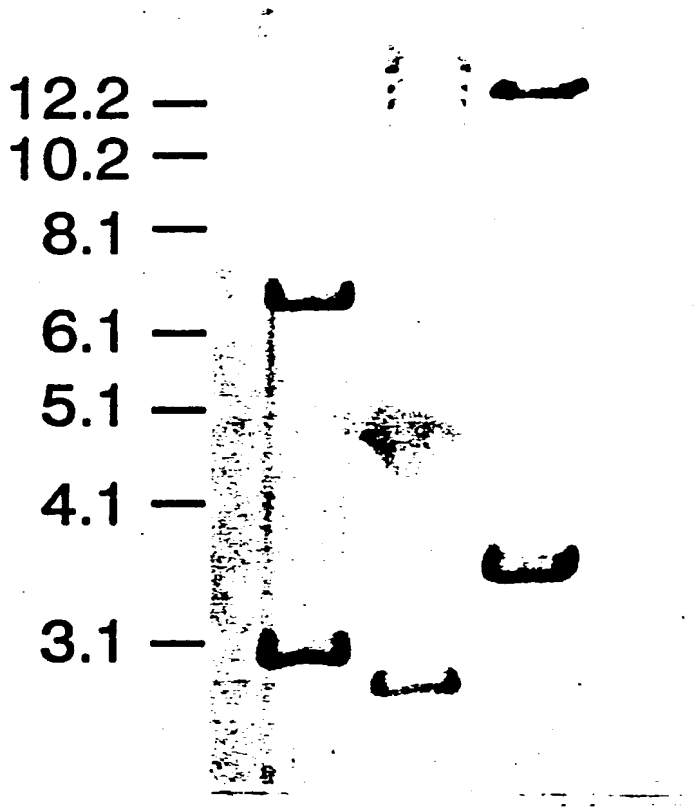
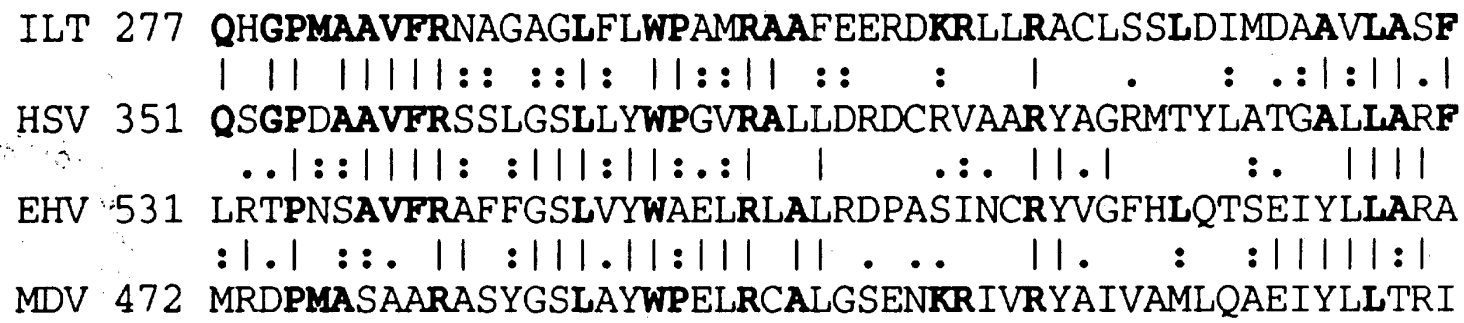


FIGURE 17


ILT 277 QHGPMAAVFRNAGAGLFLWPAMRAAFEERDKRLLRACLSSLDIMDAAVLASF
HSV 351 QSGPDAAVFRSSLGSLLYWPGVRALLDRDCRVAARYAGRMTYLATGALLARF
EHV 531 LRTPNSAVFRAFFGSLVYWAELRLALRDPASINCRYVGFHLQTSEIYLLARA
MDV 472 MRDPMASAARASYGSLAYWPELRCALGSENKRIVRYAIVAMLQAEIYLLTRI